

[illegible]

This Page Blank (uspto)

QY 254 AGCGGACGGAGAGGTGCTCATCTGAGCCAGAGAGGTTCTTACAAAGTCTTGACGGAG 313
 Db 244 AGCGGACCGGAGAGGTGCTCATCTGAGCCAGAGAGGTTCTTACAAAGTCTTGACGGAG 303
 QY 314 AGCAGAAAGCCAAAGGCTTTGAAAGGACAGTACAAATTTTGAACCATCCAGATGCTTTGATA 373
 Db 304 AGCAGAAAGCCAAAGGCTTTGAAAGGACAGTACAAATTTTGAACCATCCAGATGCTTTGATA 363
 QY 374 ATGATTTGATGACAGAGACTCTGAGAAACATCGTGAAGGGCAAAAAGGAGAGTGGCCGA 433
 Db 364 ATGATTTGATGACAGAGACTCTGAGAAACATCGTGAAGGGCAAAAAGGAGAGTGGCCGA 423
 QY 434 CCTATGATTTTGTGACACACTCAAGGTTACCAAGAGCAACGAGTGTCTTACCTCTGCGAG 493
 Db 424 CCTATGATTTTGTGACACACTCAAGGTTACCAAGAGCAACGAGTGTCTTACCTCTGCGAG 483
 QY 494 TGGTCTGTTTGAAGGGGCACTTGTGTTCTTACAGCCAGAGAGATCCGGGACATGTTCCACC 553
 Db 484 TGGTCTGTTTGAAGGGGCACTTGTGTTCTTACAGCCAGAGAGATCCGGGACATGTTCCACC 543
 QY 554 TGGGCTCTGCTGAGACACCGAGCTCCGAGCTGCTGAGAGAGTCTCCGGGAG 613
 Db 544 TGGGCTCTGCTGAGACACCGAGCTCCGAGCTGCTGAGAGAGTCTCCGGGAG 603
 QY 614 TGGCCGAGGAGGAGCCTGAGAGAGATTCTGACGAGTACACACTTCTGTGAAGCCGG 673
 Db 604 TGGCCGAGGAGGAGCCTGAGAGAGATTCTGACGAGTACACACTTCTGTGAAGCCGG 663
 QY 674 CCTTGAGAGAGTCTGCTGCTGCGGACAAAGAAAGTA-TGCCGATGTATCATCCAGAGAGA 732
 Db 664 CCTTGAGAGAGTCTGCTGCTGCGGACAAAGAAAGTA-TGCCGATGTATCATCCAGAGAGA 723
 QY 733 GTGACATATGTTGTCATCAACTGATCGTGAGAGCATTCAGAGACATTCGAAATGT 792
 Db 724 GTGACATATGTTGTCATCAACTGATCGTGAGAGCATTCGAAATGT 783
 QY 793 GACATCTGCAAAATGCGACCGAGAGAGGTCAAATGGGCGAGGCTACAGCGGACCTTTTCT 852
 Db 784 GACATCTGCAAAATGCGACCGAGAGAGGTCAAATGGGCGAGGCTACAGCGGACCTTTTCT 843
 QY 853 GAGCCAGGGGACCAACCTTGGGATGTGTACCTTGGCAAAAGGTCAATTTGGAGTCCAGC 912
 Db 844 GAGCCAGGGGACCAACCTTGGGATGTGTACCTTGGCAAAAGGTCAATTTGGAGTCCAGC 903
 QY 913 AGCAGACCCCACTGAGGGGCTGCGAGGCTCAAGGGCAGGTCTCCGGCCCGGAGATGTGT 972
 Db 904 AGCAGACCCCACTGAGGGGCTGCGAGGCTCAAGGGCAGGTCTCCGGCCCGGAGATGTGT 963
 QY 973 TCAGGAGCTGAGGCTGCGGAGCGCCACCAACCCCACTGCTTCTCTGCGGCAACCCAG 1032
 Db 964 TCAGGAGCTGAGGCTGCGGAGCGCCACCAACCCCACTGCTTCTCTGCGGCAACCCAG 1023
 QY 1033 GGGAGTGTAGCAGGAGGCTTCTTCACTCAGAGTGAATCTCAGATGTGTCACTCAG 1092
 Db 1024 GGGAGTGTAGCAGGAGGCTTCTTCACTCAGAGTGAATCTCAGATGTGTCACTCAG 1083
 QY 1093 ACTCAACTTGTGAGGACCTGACAGGGGCTTCTGAGGTTTCAAGCACTTGAAGCTCTTG 1152
 Db 1084 ACTCAACTTGTGAGGACCTGACAGGGGCTTCTGAGGTTTCAAGCACTTGAAGCTCTTG 1143
 QY 1153 CGGTTTAAAGATCCCTTAGGTCTGAGAAATGCCAGAAATGTGACAGAAAGCTTGGA 1212
 Db 1144 CGGTTTAAAGATCCCTTAGGTCTGAGAAATGCCAGAAATGTGACAGAAAGCTTGGA 1203
 QY 1213 GGCCTTGTGAGGAATGTGAGGACATTTATGGGGAATTTAGAGAGACAGCTTAGACACT 1272
 Db 1204 GGCCTTGTGAGGAATGTGAGGACATTTATGGGGAATTTAGAGAGACAGCTTAGACACT 1263
 QY 1273 GGCCTGAGCTGAGTGTGTTGTGACAGTGAACCAACAGTGGGAGAGGTTTTTCCAGTGTG 1332
 Db 1264 GGCCTGAGCTGAGTGTGTTGTGACAGTGAACCAACAGTGGGAGAGGTTTTTCCAGTGTG 1323
 QY 1333 ATCTGTTCTTACACACTCAGACATTAATCAAAAGTTTTGTGAACAAGTACTTCTT 1392

Db 1224 ATCTGTTCTTACACACTCAGACATTAATCAAAAGTTTTGTGAACAAGTACTTCTT 1383
 QY 1393 TTTTAAATGTTTAACTATGTCCTCAATGTTTCTGTTTCTGTTTCTATTAACAAGGCTGTTG 1452
 Db 1384 TTTTAAATGTTTAACTATGTCCTCAATGTTTCTGTTTCTGTTTCTATTAACAAGGCTGTTG 1443
 QY 1453 TGGCCTCAAAACCTTAAATTTCAATGACCCAGTGTGTGAGTCCAGCGTGGCCTACAGGAT 1512
 Db 1444 TGGCCTCAAAACCTTAAATTTCAATGACCCAGTGTGTGAGTCCAGCGTGGCCTACAGGAT 1503
 QY 1513 ATGGGAGCCACTGAGGAGATGTTTTCCCTTGTCTGTGTGCTTAAAGGAGAGAGCGA 1572
 Db 1504 ATGGGAGCCACTGAGGAGATGTTTTCCCTTGTCTGTGTGCTTAAAGGAGAGAGCGA 1563
 QY 1573 GGGGAGTCCCTGGAAGACCCAGACATCAACCCAGGCTTGTGCGGGCCAG 1624
 Db 1564 GGGGAGTCCCTGGAAGACCCAGACATCAACCCAGGCTTGTGCGGGCCAG 1615

RESULT 3

This Page Blank (uspto)

Accession	Sequence	Position
D8	AACCTGATGTGCACACACATCCAGGACATTTCGATATGTGTGACATCTGCAAATATGGCACCGA	813
D8	754 AACCTGATGTGCACACACATCCAGGACATTTCGATATGTGTGACATCTGCAAATATGGCACCGA	813
OY	721 GGAGGATCCAAATGGCGGAGACTCAAGACGACCTTTTCTGAGCCAGGAGACACCTCTGGG	780
D8	814 GGAGGATCCAAATGGCGGAGACTCAAGACGACCTTTTCTGAGCCAGGAGACACCTCTGGG	873
OY	781 ATACCTACCTCTGGCAAAAGGCTCACATTGGAGTCCAGACACAGACCCCACTGA	834
D8	874 ATACCTACCTCTGGCAAAAGGCTCACATTGGAGTCCAGACACAGACCCCACTGA	927

RESULT 4

```

US-09-833-381-2048
: Sequence 2048, Application US/09833381
: Patent No. US20020132090A1
:
: GENERAL INFORMATION:
: APPLICANT: Robison, Keith E.
: TITLE OF INVENTION: No. US20020132090A1 Nucleic Acid and Protein Homologs
: FILE REFERENCE: 5800-119
: CURRENT APPLICATION NUMBER: US/09/833,381
: CURRENT FILING DATE: 2001-04-11
: PRIOR APPLICATION NUMBER: 09/516,448
: PRIOR FILING DATE: 2000-02-29
: NUMBER OF SEQ ID NOS: 2050
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 2048
: LENGTH: 1648
:
: TYPE: DNA
: ORGANISM: Homo sapiens
:
: FEATURE:
: NAME/KEY: misc feature
: LOCATION: (1)...(1648)
:
: OTHER INFORMATION: n = A,T,C or G
:
: US-09-833-381-2048

```

Query Match	85.6%;	Score 714;	DB 10;	Length 1648;
Best Local Similarity	99.9%;	Pred. No. 0;		
Matches 834;	Conservative 0;	Mismatches 0;	Indels 1;	Gaps 1

QY	1	ATGGCTCTGGGGGGGAGGGGAAAGACTGGAGAGCCCGCGCCGAGGGCCGACCCGCGCAC	60
Dp	84	ATGGCTCTGGGGGGGAGGGGAAAGACTGGAGAGCCCGCGCCGAGGGCCGACCCGCGCAC	143
QY	61	CAGCGGCCCTTCTGTATAGGGGTGAGCGGCGGCACTCGACGGGGAAGTGACCGTGTGT	120
Dp	144	CAGCGGCCCTTCTGTATAGGGGTGAGCGGCGGCACTCGACGGGGAAGTGACCGTGTGT	203
QY	121	GAGAGATCATGGAATTCCTGGGACAGAACGAGGTGGAACAGCGGCGCGGGAAGTGTGTC	180
Dp	204	GAGAGATCATGGAATTCCTGGGACAGAACGAGGTGGAACAGCGGCGCGGGAAGTGTGTC	263
QY	181	ATCTGAGCGAGGACAGGTTCTACAGGTCCTGACGGCAGAGGAGAAGGCCAAGGCTTG	240
Dp	264	ATCTGAGCGAGGACAGGTTCTACAGGTCCTGACGGCAGAGGAGAAGGCCAAGGCTTG	323
QY	241	AAAGGCACTAATAATTTGACCAATCCAGATGCCCTTATATATGATTTGATGCAAGACT	300
Dp	324	AAAGGCACTAATAATTTGACCAATCCAGATGCCCTTATATATGATTTGATGCAAGACT	383
QY	301	CTGAAGACATCGTGGAGGGGCAAAACGGTGGAGGTGCGCACTATGATTTTGGACACAC	360
Dp	384	CTGAAGACATCGTGGAGGGGCAAAACGGTGGAGGTGCGCACTATGATTTTGGACACAC	443
QY	361	TCAAAGTTTACAGAACCAACGCTGTCTACCTTCGGACGCTGTTCTGTTTGAAGGCATC	420
Dp	444	TCAAAGTTTACAGAACCAACGCTGTCTACCTTCGGACGCTGTTCTGTTTGAAGGCATC	503
QY	421	TTGGTGTCTTACAGGCGAGAGATCCGGGACATGTTTCAACCTCGCGCTCTTCGAGACACC	480
Dp	504	TTGGTGTCTTACAGGCGAGAGATCCGGGACATGTTTCAACCTCGCGCTCTTCGAGACACC	563
QY	481	GACTCCGACGTCAGGCTCTTCGAAAGGTTCTCCGGACGTGCGCGCAGGGAAGGACCTG	540

This Page Blank (uspto)

614

```
DB 564 GACTCCGACGTCAAGAGTTCTCCGGGACGTCCGAGGAGGACCTG 623
QY 541 GAGCAGATTCTGACGAGTACACACCTTCGTGAAGCCGCTTCGAGAGTTCTGCTG 600
DB 624 GAGCAGATTCTGACGAGTACACACCTTCGTGAAGCCGCTTCGAGAGTTCTGCTG 683
QY 601 CCGAAGAGAGTATGCGGATGATCATCCAGAGAGTGAACAATATGTTGCTAT 659
DB 684 CCGAAGAGAGTATGCGGATGATCATCCAGAGAGTGAACAATATGTTGCTAT 743
QY 660 CAACCTGATGTCGACGACATCCAGACATTTGATGATCTGCAATATGACACCG 719
DB 744 CAACCTGATGTCGACGACATCCAGACATTTGATGATCTGCAATATGACACCG 803
QY 720 AGGAGGTCCTCAATGGCGGAGTACAGGACCTTTCTGAGCCAGGAGACCACTGG 779
DB 804 AGGAGGTCCTCAATGGCGGAGTACAGGACCTTTCTGAGCCAGGAGACCACTGG 863
QY 780 GATGCTGACCTCTGCGCAACGCTCAATTTGGAGTCCAGACAGACCCCACTGA 834
DB 864 GATGCTGACCTCTGCGCAACGCTCAATTTGGAGTCCAGACAGACCCCACTGA 918
```

```
RESULT 5
US-09-918-995-30379
Sequence 30379, Application US/09918995
Publication No. US20030073623A1
GENERAL INFORMATION:
APPLICANT: Hysq, Inc.
TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
FILE REFERENCE: 2041-756
CURRENT APPLICATION NUMBER: US/09/918,995
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: US/09/235,076
PRIORITY DATE: 1999-01-20
NUMBER OF SEQ ID NOS: 38054
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 30379
LENGTH: 472
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)..(472)
OTHER INFORMATION: n = A,T,C or G
US-09-918-995-30379
```

dele is good

```
Query Match 48.3%; Score 403; DB 11; Length 472;
Best Local Similarity 100.0%; Pred. No. 24e-207;
Matches 403; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 97 GCCAGCGGGAAGTGCACCTGTGTGAGAGATCATGAGTTGCTGGAGACGAGGTG 156
DB 41 GCCAGCGGGAAGTGCACCTGTGTGAGAGATCATGAGTTGCTGGAGACGAGGTG 100
QY 157 GACAGCGGGAAGTGCACCTGTGTGAGAGATCATGAGTTGCTGGAGACGAGGTG 216
DB 101 GACAGCGGGAAGTGCACCTGTGTGAGAGATCATGAGTTGCTGGAGACGAGGTG 160
QY 217 GCAGAGCAGAGGCGCAAGGCTTGAAGAGACAGTACATTTTGAACATCCAGATGCCCTT 276
DB 161 GCAGAGCAGAGGCGCAAGGCTTGAAGAGACAGTACATTTTGAACATCCAGATGCCCTT 220
QY 277 GATTAATATTTGATGCAACAGACTCTGAAGAACTCTGAGGGCAAAAGTGGAGGTG 336
DB 221 GATTAATATTTGATGCAACAGACTCTGAAGAACTCTGAGGGCAAAAGTGGAGGTG 280
QY 337 CCGACCTATGATTTTGGAGACACTCAAGTTACAGAGACCAAGTGTCTACCTGCG 396
DB 281 CCGACCTATGATTTTGGAGACACTCAAGTTACAGAGACCAAGTGTCTACCTGCG 340
```

This Page Blank (uspto)

```

RESULT 4
US-09-833-381-2048
; Sequence 2048, Application US/09833381
; Patent No. US20020132090A1
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: NO. US20020132090A1el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2048
; LENGTH: 1648
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1648)
; OTHER INFORMATION: n = A,T,C or G
US-09-833-381-2048

Alignment Scores:
Pred. No.: 3,648-173
Score: 1437.00
Percent Similarity: 99.64%
Best Local Similarity: 99.64%
Query Match: 99.10%
DB: 10

Length: 1648
Matches: 277
Conservative: 0
Mismatch: 0
Indels: 1
Gaps: 0

```

This Page Blank (uspto)

US-09-896-522-2 (1-277) x US-09-833-381-2048 (1-1648)

```
OY      1 MetAlaSerAlaGlyGlyGlyAspCysGlySerProAlaProGluValAlaAspArgProHis 20
DB      84 ATGGCTTCGGCGGAGCGGAGACATGCGAGACCCCGCGCGCGAGCCGACCGTCCGCAC 143
OY      21 GlnArgProPheLeuIleGlyValSerGlyGlyThrAlaSerGlyLysSerThrValCys 40
DB      144 CAGCGGCCCTTCTCTCATATGAGGCTGAGCGGCGGCACTGCCCGGAGAGTCGACCGTGT 203
OY      41 GlnLysIleMetGluLeuLeuGlyGlnAsnGluValGluGlnArgGlnArgLysValVal 60
DB      204 GAGAAAGATCATGAGTTCGTGGACAGAAACGAGTGAACAGCGGACAGCGGAGAGTGTGTC 263
OY      61 IleLeuSerGlnAspArgPheTyrLysValLeuThrAlaGluGlnLysAlaLysAlaLeu 80
DB      264 ATCCGAGCCAGGACAGAGGTTCTACAGGTCCTGACGCGCAGAGAGAGCCAGGCGCTTG 323
OY      81 LysGlyGlnTyrAsnPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThr 100
DB      324 AAGAGACAGTACATTTTGAATCCAGATGCTTGTGATGATGATGATGATGATGATGATGAT 383
OY      101 LeuLysAsnIleValGluGlyLysThrValGluValProThrTyrAspPheValThrHis 120
DB      384 CTGAAGAACATGCTGAGAGGCAAAACGGTGGAGGTCCGACCTATGATTTTGTGACACAC 443
OY      121 SerArgLeuProGluThrThrValValTyrProAlaAspValValLeuPheGluGlyIle 140
DB      444 TCMAAGTTACCAAGAGACCAAGGCTGCTACCTGCGGACGTGTCTCTGTTGAGGCGCATC 503
OY      141 LeuValPheTyrSerGlnGluIleArgAspMetPheHisIleuArgLeuPheValAspThr 160
DB      504 TTGGTGTCTACAGCCAGAGAGATCCGAGCATGTTCCACCTGCGCTTCTCTGAGACCC 563
OY      161 AspSerAspValArgLeuSerArgArgValLeuArgAspValArgArgGlyArgAspLeu 180
DB      564 GACTCCGACGTCAGCGCTGCTCGAAGAGTTCGCGGAGCGTCCGCGGAGGAGGAGCCCTG 623
OY      181 GluGlnIleLeuThrGlnTyrThrThrPheValLysProAlaPheGluGluPheCysLeu 200
DB      624 GAGCAGATTCGACGACGACGACGACGACCTTCGGAAGCCGCGCTTCGAGGAGTTCGCTG 683
OY      201 ProThrLysLysTyrAlaAspValIleIleProArgGlyValAspAsnMetValAlaIle 220
DB      684 CCGACAAAGAGATATTCGCGATGTATCATCCACAGAGAGTGAACAATATGTTGCCAT 743
OY      220 eAsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysLysThrHisArg 240
DB      744 CAACCTGATCGTGACGACATCCAGGACATTCGATGATGTATCTGCAAAATGGCACCG 803
OY      240 GGIyGlySerArgGlyArgSerTyrLysArgThrPheSerGluProGlyAspHisProGly 260
DB      804 AGGAGGGTCCCAATGGCGGAGCTACAGAGGACCTTTCTGAGCCAGGGGAGCACCCCTGG 863
OY      260 YMetLeuThrSerGlyLysArgSerHisIleuGlnIleuSerSerSerArgProHis 277
DB      864 GATGCTGACCTCTGCGAAACGTCACATTTGAGTCCAGACAGACACCCCCAC 915
```

RESULT 5

US-10-098-841-53

; Sequence 53, Application US/10098841

; Publication No. US20020197679A1

; GENERAL INFORMATION:

; APPLICANT: Tang, Y. Tom

; APPLICANT: Liu, Chenghua

; APPLICANT: Asundi, Vinod

; APPLICANT: Xu, Chongjun

; APPLICANT: Zhou, Ping

; APPLICANT: Ma, Yunding

; APPLICANT: Wang, Jian-Rui

; APPLICANT: Zhao, Qing A.

; APPLICANT: Ren, Feiyan

; APPLICANT: Ren, Rui-hong

This Page Blank (uspto)

QY 985 CTTGGGGACCCCAACCCCACTGCTTCTCTGGCCACCCAGGGAGTTAGC 1044
DB 266 CTTGGGGACCCCAACCCCACTGCTTCTCTGGCCACCCAGGGAGTTAGC 325
QY 1045 AGCGAGGCTTCTCTCACTGAGAGTGAATCTGAGATGTGTCACTGAGCTCAACTTGTCT 1104
DB 326 AGCGAGGCTTCTCTCACTGAGAGTGAATCTGAGATGTGTCACTGAGCTCAACTTGTCT 385
QY 1105 GGGACACTGACAGCGGCTTCTGAGGTTTTCAGGCACTTGGCTGCTTGAAGAT 1164
DB 386 GGGACACTGACAGCGGCTTCTGAGGTTTTCAGGCACTTGGCTGCTTGAAGAT 445
QY 1165 CCTCTAGTCACTGAGAAATGCAAGATGTGACAGAACTTGGAGGCTTCTGTAG 1224
DB 446 CCTCTAGTCACTGAGAAATGCAAGATGTGACAGAACTTGGAGGCTTCTGTAG 505
QY 1225 GAATG 1229
DB 506 GAATG 510

RESULT 8

US-09-918-995-30379
; Sequence: 30379, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO: 30379
; LENGTH: 472
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(472)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-30379

Good

Query Match 26.7%; Score 432.8; DB 11; Length 472;

Best Local Similarity 98.2%; Pred. No. 3,1e-120; Matches 437; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 177 GAGGTCGACCTTCCAGCGGAGAGTGAACCGTGTGAGAGATGATGAGTGTGG 236
DB 28 GTGGGTGAANCCAGCGGAGAGTGAACCGTGTGAGAGATGATGAGTGTGG 87
QY 237 ACAGAAAGAGTGAACAGCGGAGAGAGTGTGATCTTGAAGCAAGGTTCTA 296
DB 88 ACAGAAAGAGTGAACAGCGGAGAGAGTGTGATCTTGAAGCAAGGTTCTA 147
QY 297 CAAGGTCTGAGCAGAGCAGAAAGGCAAGGCTTGAAGAGCAGTAAATTTGACCA 356
DB 148 CAAGGTCTGAGCAGAGCAGAAAGGCAAGGCTTGAAGAGCAGTAAATTTGACCA 207
QY 357 TCCAGATGCTTTGATATGATTTGATGACAGAGCTTGAAGACATGTTGAGGCAA 416
DB 208 TCCAGATGCTTTGATATGATTTGATGACAGAGCTTGAAGACATGTTGAGGCAA 267
QY 417 AACGTGAGAGTCCGACCTATATTTTGTGACACTCAAGGTTACAGAGCACCGGT 476
DB 268 AACGTGAGAGTCCGACCTATATTTTGTGACACTCAAGGTTACAGAGCACCGGT 327
QY 477 GGTCTACCTGCGAGAGTGTCTGTTTGAAGGATCTTGTGTTTACAGCGAGAGAT 536

This Page Blank (uspto)

26:18:16:16 2003

us-09-896-522-1.rnpb

Page 7

CGTTCGCGGCGAGCGTGTCTCTGTTGAGGCGATCTTGCTGTTCTACACCGAGAGAT 387
GAGGCGATGTTCCACTCTGCGCCTCTTCGTGACACCGACTCCGACGTCAGGCTGTGCG 596
GAGGCGATGTTCCACTCTGCGCCTCTTCGTGACACCGACTCCGACGTCAGGCTGTATCG 447

592

Oy 597 AAGAGTTCTCGGAGCGTGGCCGA 621
Db 448 AAGAGTTCTCGGAGCGTGGCCGA 472

RESULT 9

Db 331 GCCCAACGGCGGC -GAGCCCTTCTTATAGCGTCAAGCGGGAACGCTTAGCGGCAG 388
Oy 202 TCGACCGTGTGTGAGAAGATCATGAGTTGCTGGACAAACGAGTGAACAGCGGCAG 261
Db 389 TCTTCGTGTGTGTAAGATCGTCAAGCTCTGGGCGAAGATGAGTGAATACTATCGCCAG 448

This page Blank (uspto)

614

```
Db      564 GACTCCGACGTCAGCTGTCTGAAAGATTCTCCGAGACGTCCGAGGAGGACCTG 623
Qy      541 GAGCAGATTCTGACCCAGTACACCACTTCTGTAAGCCGCGCTTGAGAGGATTCTGCTG 600
        624 GAGCAGATTCTGACCCAGTACACCACTTCTGTAAGCCGCGCTTGAGAGGATTCTGCTG 683
Qy      601 CCGAAGAAAGATGATGCGGATGATGATGATGATGATGATGATGATGATGATGATGAT 659
Db      684 CCGAAGAAAGATGATGCGGATGATGATGATGATGATGATGATGATGATGATGATGAT 743
Qy      660 CAACCTGATGCGGACGACATCCAGGACCTTCTGTAAGCCGCGCGCTTGAGAGGATTCTGCTG 719
Db      744 CAACCTGATGCGGACGACATCCAGGACCTTCTGTAAGCCGCGCGCTTGAGAGGATTCTGCTG 803
Qy      720 AGAAGGCTCCATGCGGACGACCTTCTGTAAGCCGCGCGCTTGAGAGGATTCTGCTG 779
Db      804 AGAAGGCTCCATGCGGACGACCTTCTGTAAGCCGCGCGCTTGAGAGGATTCTGCTG 863
Qy      780 GATGCTGACCTCTGCGCAAGGCTCACTTTGAGTCCAGGACGACCCCACTGA 834
Db      864 GATGCTGACCTCTGCGCAAGGCTCACTTTGAGTCCAGGACGACCCCACTGA 918
```

```
RESULT 5
US-09-918-995-30379
; Sequence 30379, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 30379
; LENGTH: 472
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(472)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-30379
```

date is good
102e

```
Query Match      46.3%; Score 403; DB 11; Length 472;
Best Local Similarity 100.0%; Pred. No. 2,4e-207;
Matches 403; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      97 GCCAGCGGGAAGTGCACCGTGTGAGAAGATCATGAGTTGCTGGACAGAAAGAGTG 156
Db      41 GCCAGCGGGAAGTGCACCGTGTGAGAAGATCATGAGTTGCTGGACAGAAAGAGTG 100
Qy      157 GAACAGCGGCGAGCGGAGGTGTCATCTGAGCCAGGACAGGTTCTAAGGTCTTGACG 216
Db      101 GAACAGCGGCGAGCGGAGGTGTCATCTGAGCCAGGACAGGTTCTAAGGTCTTGACG 160
Qy      217 GCAAGGCAAGAGCCCAAGGCTTGAAGAGACGTAATTTGACATCCAGATGCTTT 276
Db      161 GCAAGGCAAGAGCCCAAGGCTTGAAGAGACGTAATTTGACATCCAGATGCTTT 220
Qy      277 GATAATGATTGATGACAGGACTCTGAAGAACATCGTGAGGCAAAACGCTGAGGTG 336
Db      221 GATAATGATTGATGACAGGACTCTGAAGAACATCGTGAGGCAAAACGCTGAGGTG 280
Qy      337 CCGACCTATGATTGTTGTGACACATCAGGTTACAGAGACCAAGGTGTTACCTTCG 396
Db      281 CCGACCTATGATTGTTGTGACACATCAGGTTACAGAGACCAAGGTGTTACCTTCG 340
```

```
Qy      397 GACGTGTTCTGTTTGAAGGCACTTTGATGTTCTACAGCCAGGAGATCCGGACATGTTTC 456
Db      341 GACGTGTTCTGTTTGAAGGCACTTTGATGTTCTACAGCCAGGAGATCCGGACATGTTTC 400
Qy      457 CACCTGCGCTCTTCTGTTGACACCGACTCCGACGTGAGGCTGT 499
Db      401 CACCTGCGCTCTTCTGTTGACACCGACTCCGACGTGAGGCTGT 443
```

This Page Blank (uspto)


```

QY 141 LeuValPheTyrSerGlnGluIleArgAspMetPheHisLeuArgLeuPheValAspThr 160
DB 504 TTGGGTGTTTCAAGCAGGAGATCCGGGACATGTTCCACTCGCCTCTCTCGACAC 563
QY 161 AppSerAspValArgLeuSerArgValIleuArgAspValArgArgGlyArgAspLeu 180
DB 564 GACTCGACGTCAGGCTGTCTCGAAGAGTTCTCCGGAGCTGCGCGAGGAGGACCTG 623
QY 181 GlnGlnIleLeuThrGlnTyrThrPheValIleProAlaPheGlnGluPheCysLeu 200
DB 624 GACGCAATTTCAAGCAGACACCACTTCGTGAAGCCGCGCTTGAGAGTTCCTG 683
QY 201 ProThrTyrSerTyr-AlaAspValIleIleProArgGlyValAspAspMetValAla 220
DB 684 CCGACAAAGAAATATGCGGATGATGATCCACGAGAGTGAGCAATATGTTGCCAT 743
QY 220 eAsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysTyrTrpHisArg 240
DB 744 CAACCTGATCGTCAGCAGCATCATCAGGACATCTGAAATGTGACATCTGCAATGACACCG 803
QY 240 GGIyGlySerAsnGlyArgSerTyrTyrArgThrPheSerGluProGlyAspHisProGly 260
DB 804 AGGAGGGTTCATGCGGCGGAGCTCAAGCGGACCTTTTCTGAGCCAGGGGACCACTTGG 863
QY 260 YMetLeuThrSerGlyLysArgSerHisLeuGluSerSerSerArgProHis 277
DB 864 GATGCTGACCTCTGCGAAACGGTCACTTGGAGTCCAGCAGACAGCCCCAC 915

```

RESULT 5

```

US-09-918-995-30379
; Sequence 30379, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION.
; APPLICANT: (HySeq, Inc)
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; NUMBER OF SEQ ID NOS: 1999-01-20
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 30379
; LENGTH: 472
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(472)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-30379

```

Alignment Scores:

```

Pred. No.: 5,166-134 Length: 472
Score: 134.00 Matches: 134
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 48.38% Indels: 0
DB: 11 Gaps: 0

```

US-09-896-522-2 (1-277) x US-09-918-995-30379 (1-472)

```

QY 33 AlaSerGlyLysSerThrValCysGlnIleMetGluLeuLeuGlyGlnAsnGluVal 52
DB 41 GCCAGCGGGAAGTCGACCGTGTGTAGAAAGATCATGATGCTGCGACAGAAAGAGTGC 100
QY 53 GlnGlnArgGlnArgLysValValIleLeuSerGlnAspArgPheTyrLysValLeuThr 72
DB 101 GAACAGCGGCGACCGGAAAGGTGTCTGATCTGAGCCAGACAGGTTCTACAAAGTCTGACG 160
QY 73 AlaGlnGlnLysAlaLysAlaLeuLysGlyGlnTyrAsnPheAspHisProAspAlaPhe 92

```

```

DB 161 GCAGAGCAGAAAGGCCAAGCGCTTGAAAGACAGTACAAATTTGACCATCCAGATGCTTT 220
QY 93 AspAsnAspLeuMetHisArgThrIleLeuLysAsnIleValGlnGlyLysThrValGluVal 112
DB 221 GATTAATGATTTGATGACAGGAGCTGTGAAGAACATCTGTGAGGGCAAAACGGTGAAGGTG 280
QY 113 ProThrTyrAspPheValThrHisSerArgLeuProGluThrThrValValTyrProAla 132
DB 281 CCGACCTATGATTTTGTGACACACTCAAGTTACAGAGACACAGCGGTGCTTACCTGCG 340
QY 133 AspValValLeuPheGlnGlyIleLeuValPheTyrSerGlnGluIleArgAspMetPhe 152
DB 341 GACGTGGTCTGTTTGAAGGACATCTTGTTGTTCTACAGCAGGAGATCCGGGACATGTTT 400
QY 153 HisLeuArgLeuPheValAspThrAspSerAspValArgLeu 166
DB 401 CACCTGGCCTCTTGTGTGACACCGACTCCGAGTCAAGGCTG 442

```

This Page Blank (uspto)

This Page Blank (uspto)

QY 149 CGACACGCGGCCCTTCTGATAGGGGTGAGCGCGGCACTGCCAGCGGAAGTCGACCG 208
DB 150 CGACACGCGGCCCTTCTGATAGGGGTGAGCGCGGCACTGCCAGCGGAAGTCGACCG 209
QY 209 TGTGTGAGAGATCATGTGAGTTCGGGACAGAAAGAGTGGAAACAGCGGCAGCGGAAGG 268
DB 210 TGTGTGAGAGATCATGTGAGTTCGGGACAGAAAGAGTGGAAACAGCGGCAGCGGAAGG 269
QY 269 TGTGTGAGAGATCATGTGAGTTCGGGACAGAAAGAGTGGAAACAGCGGCAGCGGAAGG 328
DB 270 TGTGTGAGAGATCATGTGAGTTCGGGACAGAAAGAGTGGAAACAGCGGCAGCGGAAGG 329
QY 329 CCTTGAAGAGATCATGTGAGTTCGGGACAGAAAGAGTGGAAACAGCGGCAGCGGAAGG 388
DB 330 CCTTGAAGAGATCATGTGAGTTCGGGACAGAAAGAGTGGAAACAGCGGCAGCGGAAGG 389
QY 389 GGAATCTGAGAAACATCTGAGAGGCAAAACGCTGAGAGTGCAGACTATGATTTGTGA 448
DB 390 GGAATCTGAGAAACATCTGAGAGGCAAAACGCTGAGAGTGCAGACTATGATTTGTGA 449
QY 449 CACACTAGAGTTACAGAGACCAAGTGTCTACCTCGGAGCTGTCTCTTTGAGG 508
DB 450 CACACTAGAGTTACAGAGACCAAGTGTCTACCTCGGAGCTGTCTCTTTGAGG 509
QY 509 GCATCTTGGTGTCTACAGCCAGAGATCCGGGACATGTTCCACTGCGCCTTCTCTGCG 568
DB 510 GCATCTTGGTGTCTACAGCCAGAGATCCGGGACATGTTCCACTGCGCCTTCTCTGCG 569
QY 569 ACACCGAGTCCGAGCGTCAAGGCTGTCTCGAAGAGTCTCCGGAAGTGTCTCTGCGGAGG 628
DB 570 ACACCGAGTCCGAGCGTCAAGGCTGTCTCGAAGAGTCTCCGGAAGTGTCTCTGCGGAGG 629
QY 629 ACCTGAGAGAGATTTGACGCAAGTACACCACTTCTGTAAAGCCGCTTCGAGGAGTTCT 688
DB 630 ACCTGAGAGAGATTTGACGCAAGTACACCACTTCTGTAAAGCCGCTTCGAGGAGTTCT 689
QY 689 GCCTGCGGACAAAGAGTATGCGGATGATCATCCGAGAGTGGACAAATATGATTG 748
DB 690 GCCTGCGGACAAAGAGTATGCGGATGATCATCCGAGAGTGGACAAATATGATTG 749
QY 749 CCATCAACTGATGTGACAGACATCCAGAGATTCGTGAATGTGACATCTGCAATGCG 808
DB 750 CCATCAACTGATGTGACAGACATCCAGAGATTCGTGAATGTGACATCTGCAATGCG 809
QY 809 ACCGAGAGGGTCCAAATGGGCGGAGCTACAGGCGACCTTTCTGAGCAAGGAGACACC 868
DB 810 ACCGAGAGGGTCCAAATGGGCGGAGCTACAGGCGACCTTTCTGAGCAAGGAGACACC 869
QY 869 CTGGAGATGCTGACCTCTGCGAAAGGTCAATTTGGAGTCAGACAGACCACTGAG 928
DB 870 CTGGAGATGCTGACCTCTGCGAAAGGTCAATTTGGAGTCAGACAGACCACTGAG 929
QY 929 GGGCTGCGGAGCTCAAGGCGAGGTCTCCGCGCGGAGTGTGTTCAAGGAGCTGAGCCTG 988
DB 930 GGGCTGCGGAGCTCAAGGCGAGGTCTCCGCGCGGAGTGTGTTCAAGGAGCTGAGCCTG 989
QY 989 GGGAGCGCTCACCAACCACTGCT 1013
DB 990 GGGAGCGCTCACCAACCACTGCT 1014

0100

RESULT 4
US-09-764-877-2804
; Sequence 2804, Application US/09764877
; Parent No. US20020147140A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005
; CURRENT APPLICATION NUMBER: US/09/764,877
; FILING DATE: 2001-01-17
; Application data removed - refer to PALM or file wrapper

This Page Blank (uspto)

PRIOR FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 1104
SOFTWARE: pc_fl_genes Version 1.0
SEQ ID NO 546
LENGTH: 1288
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (95) .. (928)
US-10-037-270-546

this date is good

Query Match 100.0%; Score 834; DB 14; Length 1288;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 834; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	ATGCTTGGGCGGAGCGGAGAGCTGCGAGACCCCGCGCGGAGGCGGACCTCCGAC	60
DB	95	ATGCTTGGGCGGAGCGGAGAGCTGCGAGACCCCGCGCGGAGGCGGACCTCCGAC	154
QY	61	CAGCGCCCTTCTGTATAGGGGTGAGCGCGGCACTGCCAGCGGAACTGACCTGTGT	120
DB	155	CAGCGCCCTTCTGTATAGGGGTGAGCGCGGCACTGCCAGCGGAACTGACCTGTGT	214
QY	121	GAGAAATCATGAGTGTCTGCGGACGAAAGAGGTGAAACAGCGGAGCGAACTGTCT	180
DB	215	GAGAAATCATGAGTGTCTGCGGACGAAAGAGGTGAAACAGCGGAGCGAACTGTCT	274
QY	181	ATCTGAGCGAGAGCAGTCTTCAAGAGTCTGACGCGGAGAGCAGAAAGGCGAAGCTTG	240
DB	275	ATCTGAGCGAGAGCAGTCTTCAAGAGTCTGACGCGGAGAGCAGAAAGGCGAAGCTTG	334
QY	241	AAAGCAGTACATTTTTCAGATCCAGATGCTTGTATATGATTGTGACAGACT	300
DB	335	AAAGCAGTACATTTTTCAGATCCAGATGCTTGTATATGATTGTGACAGACT	334
QY	301	CTGAAGACATGCTGAGAGGCAAAAGGTGAGGTGCGGACCTATGATTTTGTACAC	360
DB	395	CTGAAGACATGCTGAGAGGCAAAAGGTGAGGTGCGGACCTATGATTTTGTACAC	420
QY	361	TCAGGTTACAGAGCAGCGGTGCTACCTGCGGAGAGTGTCTGTAGGGGATC	420
DB	455	TCAGGTTACAGAGCAGCGGTGCTACCTGCGGAGAGTGTCTGTAGGGGATC	514
QY	421	TTGCTGTTTCAAGCCAGAGAGATCCGGGACATGTTCCACTGCGCTCTTGTGACACC	480
DB	515	TTGCTGTTTCAAGCCAGAGAGATCCGGGACATGTTCCACTGCGCTCTTGTGACACC	574
QY	481	GACTCGAGTCAAGGCTGTCTGAAGAGTCTCCGGGAGTCCGCGGAGGAGGAGCTTG	540
DB	575	GACTCGAGTCAAGGCTGTCTGAAGAGTCTCCGGGAGTCCGCGGAGGAGGAGCTTG	634
QY	541	GAGCAGATTTCTGACGAGTACACCACTTCTGGAAGCGGCTTGCAGAGTTCGCTG	600
DB	635	GAGCAGATTTCTGACGAGTACACCACTTCTGGAAGCGGCTTGCAGAGTTCGCTG	694
QY	601	CCGACAAAGATATGCGGATGTATCATCCAGAGAGTGCACATATGTTGCCATC	660
DB	695	CCGACAAAGATATGCGGATGTATCATCCAGAGAGTGCACATATGTTGCCATC	754
QY	661	AACCTGATGTCAGAGATCCAGGACATTTGAAATGGGACATCGCAATGGGACCGA	720
DB	755	AACCTGATGTCAGAGATCCAGGACATTTGAAATGGGACATCGCAATGGGACCGA	814
QY	721	GAGGCTCAATGAGCGGAGCTACAGCGGACCTTTTGTGAGCAGGAGGAGGAGCTTG	780
DB	815	GAGGCTCAATGAGCGGAGCTACAGCGGACCTTTTGTGAGCAGGAGGAGGAGCTTG	874
QY	781	ATGCTGACCTTGGGAAAGGTCATATTGTGAGTCCAGAGCAGACCCCACTGA	834
DB	875	ATGCTGACCTTGGGAAAGGTCATATTGTGAGTCCAGAGCAGACCCCACTGA	928

RESULT 2
US-10-037-270-546
Sequence 546, Application US/10037270
Publication No. US20030104529A1
GENERAL INFORMATION
APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chenghua
APPLICANT: Asundi, Vinod
APPLICANT: Zhang, Jie
APPLICANT: Ren, Feiyan
APPLICANT: Chen, Rui-hong
APPLICANT: Zhao, Qing A.
APPLICANT: Wehrman, Tom
APPLICANT: Xue, Aidong J.
APPLICANT: Yang, Yonghong
APPLICANT: Wang, Jian-Rui
APPLICANT: Zhou, Ping
APPLICANT: Ma, Yungang
APPLICANT: Wang, Duntui
APPLICANT: Wang, Zhiwei
APPLICANT: Tillinghast, John
APPLICANT: Dmanac, Radoje T.
TITLE OF INVENTION: No. US20030104529A1 Nucleic Acids and
FILE REFERENCE: 784CIP2B
CURRENT APPLICATION NUMBER: US/10/037,270
CURRENT FILING DATE: 2002-01-04
PRIOR APPLICATION NUMBER: 09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/488,725

This Page Blank (uspto)

APPLICANT: Dymnac, Radoje T.
TITLE OF INVENTION: No. 6569662el Nucleic Acids and
FILE OF INVENTION: Polypeptides
FILE REFERENCE: 784CIP2B
CURRENT APPLICATION NUMBER: US/09/620,312D
CURRENT FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: 09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/488,725
PRIOR FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 1105
SOFTWARE: pc_Fl_genes Version 1.0
SEQ ID NO 546
LENGTH: 1288
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (95)..(928)
US-09-620-312D-546

Alignment Scores:
Pred. No.: 9.34e-187 Length: 1288
Score: 1450.00 Matches: 277
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-896-522-2 (1-277) x US-09-620-312D-546 (1-1288)

QY 1 MetAlaSerAlaGlyGlyGluAspCysGluSerProAlaProGluAlaAspArgProHis 20
DB 95 ATGGCTTCGGGGGAGGCGAAGACTGCGAGGCCCGCGGAGCGGACCGTCCGAC 154
QY 21 GlnArgProPheLeuIleGlyValSerGlyGlyThrAlaSerGlyLysSerThrValCys 40
DB 155 CAGCGGCCCTTCCGTATGAGGGGTGAGCGGCGGCACTCCAGCGGAGTGCAGCGTGT 214
QY 41 GluLysIleMetGluLeuLeuGlyGlnAsnGluValGluGlnArgGlnArgLysValVal 1
DB 215 GAGAGAGATCATGAGTGTCTGGGACAGACGAGGTGACACGCGGACGCGAGGTGTC 1
QY 61 IleuSerGlnAspArgPheTyLysValLeuThrAlaGluGlnLysAlaValLeu 1
DB 275 ATCTGAGCCGAGCAGAGTTCATCAAGTCTCGACGCGAGGAGGAGGAGGAGGAGG 334
QY 81 LysGlyGlnTyLysAsnPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThr 100
DB 335 AAAGGACAGTCAATTTTGACATCCAGATGCCCTTGTATATGATTTGATGACAGGACT 394
QY 101 LeuLysAsnIleValGluGlyLysThrValGluValProThrTyLysPheValThrHis 120
DB 395 CTGAGAGACATCGTGGAGGCGAAGACGCTGAGAGTTCGACCTTATTTGTGACACAC 454
QY 121 SerArgLeuProGluThrThrValValTyLysProAlaAspValValLeuPheGluGlyTle 140
DB 455 TCAAGTTTACAGAGACACGAGTGTCTACCTCGGACGCTGTCTGTGGAGGACAC 514
QY 141 LeuValPheTyLysSerGlnGluIleArgAspMetPheHisLeuArgLeuPheValAspThr 160
DB 515 TTGGTGTCTACAGCAGAGATCCGGACATGTTTCCACCTGCGCTCTTGTGGACACC 574
QY 161 AspSerAspValArgLeuSerArgArgValLeuArgAspValArgArgGlyArgAspLeu 180
DB 575 GACTCCGAGCTCAGGCTGTCTCGAAGATTCTCCGGACGTCGCGGAGGAGGAGGAGG 634
QY 181 GluGlnIleLeuThrGlnTyLysThrThrPheValLysProAlaPheGluGluPheCysLeu 200
DB 635 GAGCAGATTCTGACGCACTACACACCTTCTGTAAGCGGCGCTTGAAGAGTTTGTGCTG 694
QY 201 ProThrLysLysTyLysAlaAspValIleIleProArgGlyValAlaAspMetValAlaIle 220

RESULT 2
US-09-620-312D-546

Sequence 546, Application US/09620312D
Patent No. 6569662

GENERAL INFORMATION:

APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chenghua
APPLICANT: Asundi, Vinod
APPLICANT: Zhang, Jie
APPLICANT: Ren, Feiyan
APPLICANT: Chen, Rui-hong
APPLICANT: Zhao, Qing A.
APPLICANT: Wehrman, Tom
APPLICANT: Xue, Aildong J.
APPLICANT: Yang, Yonghong
APPLICANT: Wang, Jian-Rui
APPLICANT: Zhou, Ping
APPLICANT: Ma, Yungding
APPLICANT: Wang, Dunrui
APPLICANT: Wang, Zhiwei
APPLICANT: John Tillinghast

This Page Blank (uspto)

This Page Blank (uspto)

No.	Score	Match	Length	DB	ID	Description
1	1450	100.0	834	4	US-09-536-647-1	Sequence 1, Appl.
2	1450	100.0	1388	4	US-09-620-312D-546	Sequence 546, App
3	476	32.8	5687	4	US-09-221-017B-368	Sequence 368, App
4	451	31.1	651	4	US-09-134-001C-781	Sequence 781, App
5	434	29.9	633	4	US-09-107-532A-7370	Sequence 170, App
6	412	28.4	1230025	4	US-09-199-452A-1	Sequence 1, Appl
7	384	26.5	1830121	4	US-09-557-884-1	Sequence 1, Appl
8	384	25.6	1830121	4	US-09-557-884-1	Sequence 1, Appl
9	370.5	25.6	10011	4	US-08-961-527-76	Sequence 76, Appl
10	351.5	24.6	580073	3	US-08-545-528D-1	Sequence 1, Appl
11	156.5	10.8	4403765	3	US-09-103-840A-2	Sequence 2, Appl
12	156.5	10.8	4411529	3	US-09-103-840A-1	Sequence 1, Appl

C	13	137.5	9.5	1830121	4	US-09-557-884-1	Sequence 1, Appl
C	14	137.5	9.5	1830121	4	US-09-663-990A-1	Sequence 1, Appl
C	15	122	8.4	973	4	US-09-221-017B-536	Sequence 536, App
C	16	117	8.1	401	4	US-09-702-705-1476	Sequence 1476, App
C	17	117	8.1	401	4	US-09-736-457-1476	Sequence 1476, App
C	18	106.5	7.3	1230025	4	US-09-188-442A-1	Sequence 1, Appl
C	19	106.5	7.0	272	4	US-09-313-294A-128	Sequence 128, App
C	20	99.5	6.9	627	4	US-09-134-001C-948	Sequence 948, App
C	21	94	6.5	1554	4	US-09-252-991A-5678	Sequence 5678, App
C	22	94	6.5	1523	4	US-09-252-991A-5648	Sequence 5648, App
C	23	94	6.5	1209	4	US-09-252-991A-5548	Sequence 5548, App
C	24	93.5	6.4	591	4	US-09-252-991A-5859	Sequence 5859, App
C	25	93.5	6.4	651	4	US-09-252-991A-5948	Sequence 9948, App
C	26	93	6.4	10240	4	US-08-961-527-32	Sequence 32, Appl
C	27	92	6.3	296	4	US-09-313-294A-6207	Sequence 6207, App
C	28	92	6.3	1221	4	US-09-620-312D-162	Sequence 162, App
C	29	91.5	6.3	4128	3	US-08-793-62A-2	Sequence 2, Appl
C	30	91.5	6.3	4131	1	US-08-420-235B-2	Sequence 2, Appl
C	31	91.5	6.3	4131	5	PCT-US95-10194-2	Sequence 2, Appl
C	32	91.5	6.3	20710	1	US-08-420-235B-1	Sequence 1, Appl
C	33	91.5	6.3	20710	3	US-08-793-62A-1	Sequence 1, Appl
C	34	91.5	6.3	20710	5	PCT-US95-10194-1	Sequence 1, Appl
C	35	91.5	6.3	35100	2	US-08-770-379-18	Sequence 18, Appl
C	36	91.5	6.3	35100	3	US-08-757-669A-18	Sequence 18, Appl
C	37	91.5	6.3	35100	4	US-09-230-371A-18	Sequence 18, Appl
C	38	91	6.3	2046	4	US-09-328-352-1540	Sequence 1540, Appl
C	39	90	858	4	US-09-334-938-12	Sequence 12, Appl	
C	40	89.5	6.2	798	4	US-09-252-991A-4797	Sequence 4797, Appl
C	41	89.5	6.2	4412	4	US-09-252-991A-4748	Sequence 4748, App
C	42	89.5	6.2	4412	4	US-08-345-101A-1	Sequence 4677, App
C	43	88.5	6.1	9404	4	US-08-345-101A-1	Sequence 1, Appl
C	44	88.5	6.1	9404	4	US-09-183-688-1	Sequence 1, Appl
C	45	88.5	6.1	9404	4	US-09-515-489-1	Sequence 1, Appl

ALIGNMENTS

```

RESULT 1
US-09-536-647-1
: Sequence 1, Application US/09536647
: Patent No. 6579708
: GENERAL INFORMATION:
: APPLICANT: Ho, Yen Sen
: APPLICANT: Chunshen Tsai
: TITLE OF INVENTION: No. 6579708e1 Human Uridine Kinase
: FILE REFERENCE: GP50020
: CURRENT APPLICATION NUMBER, US/09/536,647
: NUMBER OF SEQ ID NOS: 3
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 1
: LENGTH: 834
: TYPE: DNA
: ORGANISM: Human
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (1)...(780)
: US-09-536-647-1

```

Alignment Scores:

Pred. No.:	4,51e-187	Length:	834
Score:	1450.00	Matches:	277
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	4	Gaps:	0

US-09-896-522-2 (1-277) x US-09-536-647-1 (1-834)

Qy 1 MetAlaSerAlaGlyGlyGluAspCysGluSerProAlaProGluAlaAspArgProHis 20
|||||
Db 1 ATGCGCTTCGGCGGAGGCGGAAGACTGCGAGAGCCCCCGCGCCGAGGCCGACCGTCCGCAC 60

QY 21 GlnArgProPheLeuIleGlyValSerGlyGlyThrAlaSerGlyIleSerThrValCys 40
Db 61 CAGGGCCCTTCGATAGGGGTAGCGGCGCATCGACGGGAAGTGCACGGTGT 120
QY 41 GluValIleMetGluLeuLeuGlyGlnAsnGluValGluGlnArgGlnArgValVal 60
Db 121 GAGAGATCATGAGATTCCTGGACAGAACAGAGGTGAAACGCCGACCGGAAGTGTGC 180
QY 61 IleuSerGlnAspArgPheTyrIleValIleuThrAlaGluGlnIleValAlaValIleu 80
Db 181 ATCTGAGCCAGCAGAGTTCCTCAAGGTCTGACGGCAGAGCAGAAAGCCCTTG 240
QY 81 LysGlyGlnTyrAsnPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThr 100
Db 241 AAAGAGACATCAATTTTGACATCCAGATGCCCTTTGCAATGATTTGATGACAGACT 300
QY 101 LeuValAsnIleValIleGluGlyIleValGluValProThrTyrAspPheValThrHis 120
Db 301 CTGAGAAACATCGTGGAGGCGAAACGGTGGAGGTGCCGACTATGATTTTGTGACACAC 360
QY 121 SerArgLeuProGluThrThrValIleTyrProAlaAspValIleuPheGluGlyIle 140
Db 361 TCAGGTTCACAGAGACACACGGTGTCTACCTCGGACGTGTCTGTTGAGGCGATC 420
QY 141 LeuValPheTyrSerGlnGluIleArgAspMetPheHisLeuArgLeuPheValAspThr 160
Db 421 TTGGTGTTCACACCGAGAGATCCGGACATGTTCCACCTGCCCTCTTCGTGAGAACCC 480
QY 161 AspSerAspValArgLeuSerArgValIleuArgAspValArgGlyArgAspLeu 180
Db 481 GACTCCGACGTCAAGCTGTCTCGAAGAGTTCTCCGGACGTCGCCCGAGGGAGGACTG 540
QY 181 GluGlnIleLeuThrGlnTyrThrThrPheValIleAspProAlaPheGluGluPheCysLeu 200
Db 541 GAGCAGATTCCTGACGCGATCACCACTTCCTGGAAGCCGCGCTTCGAGAGTTCGCTG 600
QY 201 ProThrIleValTyrAlaAspValIleleProArgValIleAspAsnMetValAlaIle 220
Db 601 CCGCAAGAAAGATGCGCATGTATCATCCAGAGAGTGAACAATATGTTGCCATC 660
QY 221 AsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysIleTyrHisArg 240
Db 661 AACGTATCGTGCAGCATCATCAGAGCATTCGTGATGTGACATCTGGAATGACACGA 720
QY 241 GlyGlySerAsnGlyArgSerTyrIleArgThrPheSerGluProGlyAspHisProGly 260
Db 721 GGAAGGTTCATGAGCGGAGCTACAGAGCGGACCTTTCTGAGCCAGGGGACCACTGCG 780
QY 261 MetLeuThrSerGlyIleArgSerHisIleuGluSerSerArgProHis 277
Db 781 ATGCTGACCTTGCAACGGTCAATTTGAGTCCAGACAGACCCAC 831

RESULT 2
US-09-620-312D-546
Sequence 546, Application US/09620312D
GENERAL INFORMATION
APPLICANT: Pang, Y. Tom
APPLICANT: Dir, Chingma
APPLICANT: Asundi, Vinod
APPLICANT: Zhang, Jie
APPLICANT: Ren, Feiyun
APPLICANT: Chen, Rui-hong
APPLICANT: Zhao, Qing A.
APPLICANT: Wehrman, Tom
APPLICANT: Xue, Aidong J.
APPLICANT: Wang, Yonghong
APPLICANT: Wang, Jian-Rui
APPLICANT: Zhou, Ping
APPLICANT: Ma, Yungding
APPLICANT: Wang, Dunrui
APPLICANT: Wang, Zhiwei
APPLICANT: John Tillinphast

APPLICANT: Drmanac, Radcoje T.
TITLE OF INVENTION: No. 6569662el Nucleic Acids and
FILE REFERENCE: Polypeptides
CURRENT APPLICATION NUMBER: US/09/620,312D
CURRENT FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: 09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/488,725
PRIOR FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 1105
SOFTWARE: pc_Fl_genes Version 1.0
SEQ ID NO 546
LENGTH: 1288
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (95)..(928)
US-09-620-312D-546
Alignment Scores:
Pred. No.: 9,34e-187 Length: 1288
Score: 1450.00 Matches: 277
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0
US-09-896-522-2 (1-277) x US-09-620-312D-546 (1-1288)
QY 1 MetAlaSerAlaGlyGlyIleAspCysGluSerProAlaProGluAlaAspArgProHis 20
Db 95 ATGCTTCGGCGGAGAGCGAAGCTCGAAGGCCCGCGCGGAGCCGACCGTCCGAC 154
QY 21 GlnArgProPheLeuIleGlyValSerGlyIleThrAlaSerGlyIleSerThrValCys 40
Db 155 CAGCGCCCTTCCTGATAGGGGTGACCGCGGACCTCCAGCGGAGAGTGCACGGTGT 214
QY 41 GluValIleMetGluLeuLeuGlyGlnAsnGluValGluGlnArgGlnArgValVal 60
Db 215 GAGAGATCATGAGATTCCTGGACAGAAAGAGTGAACAGCGCGAGCGGAGAGTGTGC 274
QY 61 IleuSerGlnAspArgPheTyrIleValIleuThrAlaGluGlnIleValAlaValIleu 80
Db 275 ATCTGAGCCAGAGCAGTTCCTCAAGGTCTTGAAGCGCAGAGCAAGGCGCAAGGCTTG 334
QY 81 LysGlyGlnTyrAsnPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThr 100
Db 335 AAAGACAGTCAATTTTGAACATCCAGATGCTTTGATATGATTTGATGACAGAGACT 394
QY 101 LeuValAsnIleValIleGluGlyIleThrValGluValProThrTyrAspPheValThrHis 120
Db 395 CTGAGAAACATCGTGGAGGCGCAAAACGGTGAAGTCCCGCATGATTTTGTGACACAC 454
QY 121 SerArgLeuProGluThrThrValIleTyrProAlaAspValIleuPheGluGlyIle 140
Db 455 TCAGGTTCACAGAGACACAGGTGTCTACCTCGGACGATGTTCTGTTGGGGCATC 514
QY 141 LeuValPheTyrSerGlnGluIleArgAspMetPheHisIleuArgLeuPheValAspThr 160
Db 515 TTGGTGTTCACAGAGACAGAGATCCGGACATTTCCACTGGGCTCTTCGAGACACC 574
QY 161 AspSerAspValArgLeuSerArgValIleuArgAspValArgArgGlyArgAspLeu 180
Db 575 GACTCCGAGCTCAGGCTGTCTCGAAGGTCTCCGGAGCGTGGCGGAGGAGGAGACTG 634
QY 181 GluGlnIleLeuThrGlnTyrThrThrPheValIleProAlaPheGluGluPheCysLeu 200
Db 635 GAGCAGATTCGACGAGTACACACACTTCGTGAACCGGCTTCGAGAGTTCGCTG 694
QY 201 ProThrIleValTyrAlaAspValIleleProArgValIleAspAsnMetValAlaIle 220

Db 695 CCGACAAAGAGATGCGATGATCATCCAGAGAGTGCACATATGTGCCATC 754
Qy 221 AsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysIleValTyrPheArg 240
Db 755 AACCTGATCGTGCAGCATCCAGACATCTGATAGTGAACATCTGCAATGGCACCGA 814
Qy 241 GlyGlySerAsnGlyArgSerTyrIleValArgThrPheSerGluProGlyAspHisProGly 260
Db 815 GAGGGGTCCATGGGCGAGACTCAAGCGGACCTTTCTTAGCCAGGGGACCACTGGG 874
Qy 261 MetLeuThrSerGlyIleValArgSerHisLeuGluSerSerArgProHis 277
Db 875 ATGCTGACCTGTGCAAAACGTCATCTTGAGTCCAGCAGACCCAC 925
RESULT 3
US-09-221-017B-368
Sequence 368, Application US/09221017B
Patent No. 6444799
GENERAL INFORMATION:
APPLICANT: Ross, Bruce C.
TITLE OF INVENTION: P. GINGIVALIS NUCLEOTIDES AND USES THEREOF
NUMBER OF SEQUENCES: 1120
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER
STREET: 755 PAGE MILL ROAD
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304-1018
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows
SOFTWARE: FASTSEQ for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/221,017B
FILING DATE: 23-DEC-1998
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: PP1182
FILING DATE: 31-DEC-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PP1546
FILING DATE: 30-JAN-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PP2911
FILING DATE: 09-APR-1998
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: PCT/AU98/01023
FILING DATE: 10-DEC-1998
ATTORNEY/AGENT INFORMATION:
NAME: MONROY, Gladys H
REGISTRATION NUMBER: 32,430
REFERENCE/DOCKET NUMBER: 27340-20021.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-813-5600
TELEFAX: 650-494-0792
TELEX: 706141
INFORMATION FOR SEQ ID NO: 368:
SEQUENCE CHARACTERISTICS:
LENGTH: 5687 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULAR TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: UNKNOWN
ORIGINAL SOURCE:
ORGANISM: PORYPHYROMONAS GINGIVALIS
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1...5687
US-09-221-017B-368

Alignment Scores:
Pred. No.: 2,276-53 Length: 5687
Score: 476.00 Matches: 91
Percent Similarity: 70.14 Conservative: 57
Best Local Similarity: 43.13 Mismatches: 47
Query Match: 32.83 Indels: 16
DB: 4 Gaps: 4
US-09-896-522-2 (1-277) x US-09-221-017B-368 (1-5687)
Qy 25 LeuIleGlyValSerGlyIleThrAlaSerGlyIleSerThrValCysGluIleMet 44
Db 4706 ATATAGAGATGACAGCGGAGAGTGTCTCGGCAAGATACATGTGTGAAAAAGCTCTGT 4765
Qy 45 GluLeuLeuGlyGlnAsnGluValGluGlnArgGlnArgValIleLeuSerGln 64
Db 4766 GAGGCTTGTGTCGAGAGAGAT-----GTGCTGTACTCTGTCTAC 4804
Qy 65 AspArgPheTyrIleVal-----IleThrAlaGluGlnIleValAlaLeuIle 81
Db 4805 GATTACTATTACAGGCGCAATGACACCTCTCTTGAGGAAAGAAAG----- 4855
Qy 82 GlyGlnTyrAsnPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThrLeu 101
Db 4856 -----CTGACTACGACCATCCCATGCTTTCGACACGATATGTCGACGATATT 4909
Qy 102 LysAsnIleValGluGlyIleThrValGluValProThrTyrAspPheValThrHisSer 121
Db 4910 CTCTCTCTGAGGAGGCAAGCAAGATAGAGCGTCCGCTATATCTTCTAGAGCACAAT 4969
Qy 122 ArgLeuProGluThrThrValIleTyrProAlaAspValIleLeuPheGluGlyIleLeu 141
Db 4970 CGTTGCAAGAAAGTAACCTGTCTGCTGCAAGATGATGATGATGATGATGATGATGATG 5029
Qy 142 ValPheTyrSerGlnGluIleArgAspMetPheIleLeuArgPheValAspThrAsp 161
Db 5030 ATATTGAAAGCAAGAGGCGGATCGATGATGATGATGATGATGATGATGATGATGATGATG 5089
Qy 162 SerAspValArgLeuSerArgArgValLeuArgAspValArg--ArgGlyArgAspLeu 180
Db 5090 GCGGATATTGATGATGCGCGCGCTTGCGCGCATGTCCAGAAACGGGAGCAGCAATATAG 5149
Qy 181 GluGlnIleLeuThrGlnTyrThrPheValIleProAlaPheGluGluIleCysLeu 200
Db 5150 GATCGGATTTGACACAAATCACTTCAGTACGTTGCGCTATGACACGAGATTTTGTGAA 5209
Qy 201 ProThrIleValTyrAlaAspValIleIleProArgGlyValAspAsnMetValAlaIle 220
Db 5210 CATCCAGCGGTAGCGCGCATCTGATCAATCCGGAAGTGGATTCAATTGGTGGCGCTC 5269
Qy 221 AsnLeuIleValGlnHisIleGlnAspIleLeu 231
Db 5270 TCATCTCTGTGCAAAAAATCCGATCGGTGATT 5302
RESULT 4
US-09-134-001C-781
Sequence 781, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
PRIOR FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 781
LENGTH: 651

```

1 TYPE: DNA
2 ORGANISM: Staphylococcus epidermidis
3 US-09-134-001C-781
4
5 Alignment Scores:
6 Pred. No.: 1,5e-51 Length: 651
7 Score: 451.00 Matches: 91
8 Percent Similarity: 63.68% Conservative: 44
9 Best Local Similarity: 42.92% Mismatch: 61
10 Query Match: 31.10% Indels: 16
11 DB: 4 Gaps: 4
12
13 US-09-896-522-2 (1-277) x US-09-134-001C-781 (1-651)
14
15 QY 25 LeuilegylValSerlgylYthAlaserGlySerThrValCySGluysIleMet 44
16   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
17 Db 43 ATTTATGGCATTTGACAGGGGGTTCCAGGCTCTGGAAACACACAGTTACTAATGCAATATG 102
18
19 QY 45 GluIeuLeuGlyIuAenGluValGluGlnArgGlnArgLysValValIleLeuSerGln 64
20   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
21 Db 103 AAGAATTATTA-----GAAGTCATTAAGTTCGACCTTTAGCACAA 141
22
23 QY 65 AsparGpPheTyrLys-----ValIeuThrIaGluGlnLysAlaLeuLys 81
24   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
25 Db 142 GATTATTAATCTAATAAGATCAATCACTTAATCTTTGAAAGAACGGCTTGAACA----- 195
26
27 QY 82 GlycInTyrAenPheAspHisProAspAlaPheAspAsnAbLeuMetHisArgThrLeu 101
28   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
29 Db 196 -----AATTTATGACCACCCCATTTGATTCATTTGATATGATTAATTATTCATTAATTTA 246
30
31 QY 102 LysAsnIleValGluGlyLysThrValGluValProThrTyrAspPheValThrHisSer 121
32   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
33 Db 247 AAGAGTTTAAGAAATGGAAACCGATAGAGTACCTTACGTACGATTATCTTCAACATACA 306
34
35 QY 122 ArgLeuProGluThrThrValValTyrProAlaAspValValIeuPheGluGlyIleLeu 141
36   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
37 Db 307 CGTGAAGTAAGAAACAATGATTCATTTGATTCACAAAGATGTTATTATTCGTAGAAGGTATCTT 366
38
39 QY 142 ValPheTyrSerGlnGluIleArgAspMetPheHisLeuArgLeuPheValAspThrAsp 161
40   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
41 Db 367 GCTTTAGAACCAACACATTCAGACATGATGATGAGAAATTTATGTCGATACATGAT 426
42
43 QY 162 SerAspValArgLeuSerArgArgValIeuArgAspValArg---ArgGlyArgAspLeu 180
44   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
45 Db 427 GCTGATTTACAAATTTTACGTAGGCTTCAAGAGATCTAAAGAGCGGTGCTACATG 486
46
47 QY 181 GluGlnIleLeuThrGlnTyrThrThrPheValLysProAlaPheGluGlnPheCysLeu 200
48   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
49 Db 487 GAATCAGTAATTAATCAATCAATCTTAATGTATAGACTATGACATGACATTTATTTGAA 546
50
51 QY 201 ProThrLysLysTyrAlaAspValIleIleProArgGlyValAspAsnMetValAlaIle 220
52   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
53 Db 547 CCTCAAAAAAGACATGTCGATCATCATATTTCTTGAAGAGGACAGCAATAAAGTTGCATA 606
54
55 QY 221 AsnIeuIleValGlnHisIleGlnAspIleLeuAsn 232
56   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
57 Db 607 GATATTATGACTACTAAATCCAGCTACCTCGTAAGC 642
58
59 RESULT 5
60 US-09-107-532A-1370
61 / Sequence 1370, Application US/09107532A
62 / Patent No. 6583275
63 / GENERAL INFORMATION:
64 / APPLICANT: Lynn A Doucette-Stamm and David Bush
65 / TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
66 / ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
67 / NUMBER OF SEQUENCES: 7310
68 / CORRESPONDENCE ADDRESS:
69 / ADDRESSEE: GENOME THERAPEUTICS CORPORATION
70 / STREET: 100 Beaver Street
71 / CITY: Waltham
72 / STATE: Massachusetts
73 / COUNTRY: USA

```

```

1 ZIP: 02354
2
3 COMPUTER READABLE FORM:
4 MEDIUM TYPE: CD-ROM ISO9660
5
6 COMPUTER: PC
7 OPERATING SYSTEM: <unknown>
8 SOFTWARE: ASCII
9
10 CURRENT APPLICATION DATA:
11 APPLICATION NUMBER: US/09/107,532A
12 FILING DATE: 30-Jun-1998
13
14 PRIOR APPLICATION DATA:
15 APPLICATION NUMBER: 60/085,598
16 FILING DATE: 14 May 1998
17 APPLICATION NUMBER: 60/051571
18 FILING DATE: July 2, 1997
19
20 ATTORNEY/AGENT INFORMATION:
21 NAME: Ariniello, Pamela Deneke
22 REGISTRATION NUMBER: 40,489
23 REFERENCE/DOCKET NUMBER: GTC-012
24 TELECOMMUNICATION INFORMATION:
25 TELEPHONE: (781)893-5007
26 TELEFAX: (781)893-8277
27
28 INFORMATION FOR SEQ ID NO: 1370:
29
30 SEQUENCE CHARACTERISTICS:
31 LENGTH: 633 base pairs
32 TYPE: nucleic acid
33 STRANDEDNESS: double
34 TOPOLOGY: circular
35
36 MOLECULE TYPE: DNA (genomic)
37
38 HYPOTHETICAL: NO
39
40 ANTI-SENSE: NO
41
42 ORIGINAL SOURCE:
43 ORGANISM: Enterococcus faecium
44
45 FEATURE:
46 NAME/KEY: misc_feature
47 LOCATION: (B) LOCATION 1...633
48
49 SEQUENCE DESCRIPTION: SEQ ID NO: 1370:
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970

```



```

NAME/KEY: misc.feature
LOCATION: (180001)..(195000)
OTHER INFORMATION: -na or c or g or t
NAME/KEY: misc.feature
LOCATION: (195001)..(210000)
OTHER INFORMATION: -na or c or g or t
NAME/KEY: misc.feature
LOCATION: (210001)..(225000)
OTHER INFORMATION: -na or c or g or t
NAME/KEY: misc.feature
LOCATION: (225001)..(240000)
OTHER INFORMATION: -na or c or g or t
NAME/KEY: misc.feature
LOCATION: (240001)..(255000)
OTHER INFORMATION: -na or c or g or t
NAME/KEY: misc.feature
LOCATION: (255001)..(270000)
OTHER INFORMATION: -na or c or g or t

```

```

LOCATION: (270601)..(285000)
OTHER INFORMATION: -na or c or g or t
NAME/KEY: misc_feature
LOCATION: (285001)..(300000)
OTHER INFORMATION: -na or c or g or t
NAME/KEY: misc_feature
LOCATION: (300601)..(315000)
OTHER INFORMATION: -na or c or g or t
NAME/KEY: misc_feature
LOCATION: (315001)..(330000)
OTHER INFORMATION: -na or c or g or t
NAME/KEY: misc_feature
LOCATION: (330001)..(345000)
OTHER INFORMATION: -na or c or g or t
NAME/KEY: misc_feature
LOCATION: (345001)..(360000)
OTHER INFORMATION: -na or c or g or t
NAME/KEY: misc_feature
LOCATION: (360001)..(375000)
OTHER INFORMATION: -na or c or g or t
NAME/KEY: misc_feature
LOCATION: (375001)..(390000)
OTHER INFORMATION: -na or c or g or t
NAME/KEY: misc_feature
LOCATION: (390001)..(405000)
OTHER INFORMATION: -na or c or g or t
NAME/KEY: misc_feature
LOCATION: (405001)..(420000)
OTHER INFORMATION: -na or c or g or t
NAME/KEY: misc_feature
LOCATION: (420001)..(435000)
OTHER INFORMATION: -na or c or g or t
NAME/KEY: misc_feature
LOCATION: (435001)..(450000)
OTHER INFORMATION: -na or c or g or t
NAME/KEY: misc_feature
LOCATION: (450001)..(465000)
OTHER INFORMATION: -na or c or g or t
NAME/KEY: misc_feature
LOCATION: (465001)..(480000)
OTHER INFORMATION: -na or c or g or t
NAME/KEY: misc_feature
LOCATION: (480001)..(495000)
OTHER INFORMATION: -na or c or g or t
NAME/KEY: misc_feature
LOCATION: (495001)..(510000)
OTHER INFORMATION: -na or c or g or t
NAME/KEY: misc_feature
LOCATION: (510001)..(525000)
OTHER INFORMATION: -na or c or g or t
NAME/KEY: misc_feature
LOCATION: (525001)..(540000)
OTHER INFORMATION: -na or c or g or t
NAME/KEY: misc_feature

```

LOCATION: (540001)..(555000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (555001)..(570000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (570001)..(585000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (585001)..(600000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (600001)..(615000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (615001)..(630000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (630001)..(645000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (645001)..(660000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (660001)..(675000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (675001)..(690000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (690001)..(705000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (705001)..(720000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (720001)..(735000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (735001)..(750000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (750001)..(765000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (765001)..(780000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (780001)..(795000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (795001)..(810000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (810001)..(825000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (825001)..(840000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (840001)..(855000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (855001)..(870000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (870001)..(885000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (885001)..(900000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (900001)..(915000)

OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
Alignment Scores:
Pred. No.: 9.05e-41 Length: 123025
Score: 412.00 Matches: 84
Percent Similarity: 63.01% Conservative: 54
Best Local Similarity: 38.36% Mismatches: 59
Query Match: 28.41% Indels: 22
DB: 4 Gaps: 6
US-09-896-522-2 (1-277) x US-09-198-452A-1 (1-123025)
Qy 25 Leuileg1valserg1vg1ythr1a1serg1lyleser1h1val1cysg1uyl1emet 44
Db 835771 ATTATTGAAATTTCAGAGAGGTTTGAGAGCTGGAGAAACCCCTTAACCCAAACATTTAA 835712
Qy 45 Gluleuleug1ygn1nsg1uval1gu1n1arg1n1arg1ys1val1le1se1rg1n 64
Db 835711 GAAATTTTCGTGAGAT-----CTGAGTGTTATCTGCGCA 835676
Qy 65 Asp1argp1hery1lys-----Val1leu1thr1al1gl1u1n1ys1al1aleu1ys 81
Db 835675 GATTAATTATTACAAAGATAGATCTCATTAATCTGAAAGAACGTCCAATTTAAT-- 835619
Qy 82 Gl1y1n1tr1asn1h1eas1p1s1pro1asp1al1phe1asp1asn1h1e1s1arg1h1leu 101
Db 835618 -----TGGATCTCTCGAGCGCTTTGATTAATGACTTTAATTTAGGTAAT 835511
Qy 102 Lys1asn1leval1gl1u1y1thr1val1gl1uval1pro1h1tr1asp1h1e1val1thr1s1ser 121
Db 835570 AAACGCTAAATAATATGATGATTTGCCAAGCCCAAGTTTGTATTTAGGTAAT 835511
Qy 122 Arg---Leu1pro1h1thr1h1r1h1val1y1r1pro1al1asp1val1leu1p1he1gu1y1le 140
Db 835510 CGATCTAAACGAGATAGAAACGATCTATCCATTAAGTTATTTGTTGAAGGTAAT 835451
Qy 141 Leu1val1p1hery1ser1g1n1le1arg1asp1met1phe1s1leu1arg1h1e1val1asp1thr 160
Db 835450 CTGCTCTTGAATAATCAAGACTTAAGATCTTAATGATTAATTTGATTTGACACCC 835391
Qy 161 Asp1ser1asp1val1arg1leu1ser1arg1val1leu1arg1asp1val1arg---Arg1y1arg1asp 179
Db 835390 GATGCTGATGAAGATACATACCGCTATGTTGAGATGTTCAAGAACAGAGATAGC 835331
Qy 180 Leu1gl1u1n1leu1h1rg1n1y1r1h1r1h1p1he1val1ys1pro1al1phe1gu1n1p1he1ys 199
Db 835330 GTGAGCTGATCAATGCTGTTATCTTATGATTAAGCTTAATGATGAAGAAATTTATA 835271
Qy 200 Leu1pro1h1r1h1s1y1s1tyr1al1asp1val1le1le1pro1arg1y1val1asp1n1met1val1a 219
Db 835270 GAGCCGACTCGGAATATGCTGATATCATTTGATCAATGAAATTAACGACAAACGATAGA 835211
Qy 220 Ile1asn1leu1leval1gn-----His1le1gn1asp1le1leu1nsg1y1asp 234
Db 835210 ACAATATATTTTGCACAGAAATTAATAATCAATTAGAGAAATCCCTGAAAGCCAT 835154
RESULT 7
US-09-557-884-1
Sequence 1, Application US/09557884
Patent No. 6506581
GENERAL INFORMATION:
APPLICANT: Fleischmann et al.
TITLE OF INVENTION: The Nucleotide sequence of
the Haemophilus influenzae Rd Genome, Fragments
thereof, and Uses Thereof
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSER: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
STATE: MD
COUNTRY: USA


```

Db 146027 ATCATGCCATCAGAGTGCATCTGCATCAGGCAAAAGTCTTATCGCTTCACATGTTCC 146086
Qy 41 GtlystIleMetGluLeuLeuGlyGlnAsnGluValGlnGlnArgGlnArgValVal 60
Db 146087 AAAGGCTCTTAATGAACCTGAGCGCCCAAAATTTGT----- 146125
Qy 61 IleuSerGlnAspArgPheTyrIleValLeuThrIleGluIleValAlaValAlaLeu 80
Db 146126 ATCATTTACTGAAGACAGTACTATTAAGATCAAAAGTCTTTAGAAATGACGACGGGTA 146185
Qy 81 LysGlyGlnTyrAsnPheAspPheIleProAspAlaPheAspAsnAspLeuMetHisArgThr 100
Db 146186 AAA-----ACGAAATTACGATCATCAAACTCTATGATCCGATTTACTTATCCACAT 146239
Qy 101 LeuLysAsnIleValGluGlyLysThrValGlnValProThrTyrAspPheValThrHis 120
Db 146240 TTAATAAATCTPAAAAATGAGAGTGCATGATGATGCTGTTTATGATGTAAGTGAACAT 146299
Qy 121 SerArgLeuProGluThrThrValValTyrProAlaAspValValLeuPheGluGlyIle 140
Db 146300 ACTGCGACCAACGAAACAAACCATTTTCACACCAAAACGATCTTAATTTTGAAGGATTT 146359
Qy 141 LeuValPheTyrSerGlnGluIleArgPheMetPheHisLeuArgLeuPheValAspThr 160
Db 146360 TTATTTACTTACCGATGAAAGAGTACGCGCAATTAAGCCGATATTCTGTATTCGTAGACA 146419
Qy 161 AsperAspValAlaGluSerArgArgValLeuArgAspVal---ArgArgGlyArgAsp 179
Db 146420 CCATCTGATTTGTTGTTTCATCCGCGTTTACACGATGATGAGAAACGCGTCCGCTCT 146479
Qy 180 LeuGluGlnIleLeuThrGlnTyrThrPheValLysProAlaPheGluGluPheCys 199
Db 146480 CTACATAGAGGATTTGATCATATTCGTCACACCGTCCCAATGTTCTTCAATTTATTT 146539
Qy 200 LeuProThrLysLysTyrAlaAspValIleIleProArgGlyValAspAsnMetValAla 219
Db 146540 GAGCGCTCTAAACAATGCGGATATTTGTCAATCTCTCGTGTGTAATAAATCGCATTCG 146599
Qy 220 IleAsnLeuIleValGlnHisIleGlnAspIleLeuAsn 232
Db 146600 ATCATATGTTAAAAAGCTCAAAATCTTCATTTATTTGAAT 146638

RESULT 9
US-08-961-527-76/c
Sequence 76, Application US/08961527
GENERAL INFORMATION:
APPLICANT: Charles Kunsch
TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
NUMBER OF SEQUENCES: 391
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
OPERATING SYSTEM: HP Vectra 486/33
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/961,527
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB340P1

```

```

TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 76:
SEQUENCE CHARACTERISTICS:
LENGTH: 10011 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-961-527-76

Alignment Scores:
Pred. No.: 1,25e-38 Length: 10011
Score: 370.50 Matches: 69
Percent Similarity: 66.89% Conservative: 30
Best Local Similarity: 46.62% Mismatches: 48
Query Match: 25.55% Indels: 1
Gaps: 1

US-09-896-522-2 (1-277) x US-08-961-527-76 (1-10011)
Qy 85 AsnPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThrIleLysAsnIle 104
Db 10000 AACTATGACCATCTTTGGCTTTGATACAGACTGATGATCGAGCAGATTAAAGAAATTG 9941
Qy 105 ValGluGlyLysThrValGlnValProThrTyrAspPheValThrHisSerArgLeuPro 124
Db 9940 TTGGCGAGGGGCGCGGCGGATCCGACCTGACCTATTAACAGACATACAGAGTACGAGTAGC 9881
Qy 125 GluThrThrValValTyrProAlaAspValValLeuPheGluGlyIleLeuValPheTyr 144
Db 9880 AAGACCATGCTGAGAGACCTCAAGATGCTTATATGCTTAGAGGATTTTGGCTTGAG 9821
Qy 145 SerGlnGluIleArgAspMetPheHisLeuArgLeuPheValAspThrAspSerAspVal 164
Db 9820 GACAAAGGCTGCGCGCATTTGATGATCAAAATTTTGTGATCGGATGACGATGACGATGTC 9761
Qy 165 ArgLeuSerArgArgValLeuArgAspVal---ArgArgGlyArgAspLeuGluGlnIle 183
Db 9760 CGCATTTATTTGTGATTCAAACGCGATATGAGAGACGCTGGCCGTGATACCGGTT 9701
Qy 184 LeuThrGlnTyrThrThrPheValLysProAlaPheGluGluPheCysLeuProThrLys 203
Db 9700 ATTAACCACTAGCTTAGGTGTGTCAAACCAATGTACACACGTTATTCAGCTCAATAG 9641
Qy 204 LysTyrAlaAspValIleIleProArgGlyValAspAsnMetValAlaIleAsnLeuIle 223
Db 9640 CGTTATGCTGATATTCGTCATTTCTCGAAGGGGTTAGCAATACCGTGTGCTGACCTGTTG 9581
Qy 224 ValGlnHisIleGlnAspIleLeu 231
Db 9580 ACGACCAAGATTCCAAAGATTTTG 9557

RESULT 10
US-08-545-528D-1
Sequence 1, Application US/08545528D
GENERAL INFORMATION:
APPLICANT: Fraser et al.
TITLE OF INVENTION: Nucleotide Sequence of the Mycoplasma Genitium Genome, Fragment
PATENT NO. 6537773
TITLE OF INVENTION: Thereof, and Uses Thereof
FILE REFERENCE: PB193P1
CURRENT APPLICATION NUMBER: US/08/545,528D
CURRENT FILING DATE: 1995-10-19
PRIOR APPLICATION NUMBER: US 08/488,018
PRIOR FILING DATE: 1995-06-07
PRIOR APPLICATION NUMBER: US 08/473,545
PRIOR FILING DATE: 1995-06-07
NUMBER OF SEQ ID NOS: 1
SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
LENGTH: 580073

```

```

; TYPE: DNA
; ORGANISM: Mycoplasma genitalium
; US-08-545-528D-1

Alignment Scores:
Pred. No.: 4,25e-33 Length: 580073
Score: 351.50 Matches: 72
Percent Similarity: 60.66% Conservative: 56
Best Local Similarity: 34.12% Mismatches: 70
Query Match: 24.24% Indels: 13
DB: 4 Gaps: 4

US-09-896-522-2 (1-277) x US-08-545-528D-1 (1-580073)

Qy 21 GlnArgProheLeuIleGlyValSerGlyThrAlaSerGlyLysSerThrValCys 40
Db 480697 GAAATAAGGATTTTAAAGTTGCAATTAGTGTTGATGTTCCAGGAAACCTACGTTGCT 480756
Qy 41 GlnLysIleMetGlnLeuGlyGlnAsnGluValGluGlnArgGlnArgLysValVal 60
Db 480757 GAAATGATTTATCACTTTT-----AGTAAAAAATTAAAGTTGCG 480798
Qy 61 IleLeuSerGlnAspArgPheTyLysValLeuThrAlaGluGlnLysAlaLysValLeu 80
Db 480799 ATCATCTGTCAGATTAATACTTACAG-----TCTATAATAATAAGCCATTATTA 480852
Qy 81 LysGlyGlnTyLysAsnPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThr 100
Db 480853 AGAAAAACAATAAATTGATTCATCTGATGCTTTTATGAAAACTTTTAAAGATCACAC 480912
Qy 101 LeuLysAsnIleValGluGlyLysThrValGluValProThrTyLysAspPheValThrHis 120
Db 480913 ATGGAAGATCTTCTTAAACGGTAGATAGTATGTTCTCTTATATGATCACTTATCTAT 480972
Qy 121 SerArgLeuProGluThrThrValValTyProAlaAspValValLeuPheGluGlyIle 140
Db 480973 ACCAGAGCTAAATAAACAAGAAAAATTGTCCAATGATGTTGTTATTCTAGAGGCTTA 481032
Qy 141 LeuValPheTySerGlnGluIleArgAspMetPheHisLeuArgLeuPheValAspThr 160
Db 481033 ATCCAGATGATTTGATGAAAAATTATCAAGACTTCTTAAGCTAAATAATTATTAAGAAACA 481092
Qy 161 AspSerAspValArgLeuSerArgArgValLeuArgAspValArgArgGlyArgAspLeu 180
Db 481093 AATGGGAAAGAAAGCTTTAATTAAGAAATGAAAGAAAGAAAGAAAGAAAGAAATTT 481152
Qy 181 GluGlnIleLeuThrGlnTyLysThrPheValLysProAlaPheGluGluPheCysLeu 200
Db 481153 GATTCATATTATTAACAGTACGCGGAATAGTACCAATGATGAAATATTTGTAGAA 481212
Qy 201 ProThrLysLysTyLysAlaAspValIleIlePro-----ArgGlyValAspAsn 216
Db 481213 AAAATGAAAGCAATGCTGATTAATCTGCTTGAAGTCAACGACAGAAAGAAAGTACAA 481272
Qy 217 MetValAlaIleAsnLeuIleValGlnHisIle 227
Db 481273 AGTGTA---TTGATGTGCAATTGAACACTTA 481302

RESULT 11
; US-09-103-840A-2/c
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103.840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2

```

```

; SOFTWARE: Patent Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
; US-09-103-840A-2

Alignment Scores:
Pred. No.: 3.89e-05 Length: 4403765
Score: 156.50 Matches: 70
Percent Similarity: 39.16% Conservative: 51
Best Local Similarity: 22.65% Mismatches: 127
Query Match: 10.79% Indels: 61
DB: 3 Gaps: 13

US-09-896-522-2 (1-277) x US-09-103-840A-2 (1-4403765)

Qy 3 SerAlaGlyGlyGlnAspCysGlnSerProAlaProGluAlaAspArgProHisGlnArg 22
Db 1219499 GCGGCCACCGCGGAATTCTCGCGCAGACCCCAACAAACCCGACCGCGCG-----GTG 1219446
Qy 23 ProPheLeuIleGlyValSerGlyThrAlaSerGlyLysSerThrValCysGluLys 42
Db 1219445 CCGTTATCATCTGCTGTGTGGCGGCACTGTGGCGGTGGCGCAAGTCCACACCGCCCGCTG 1219386
Qy 43 IleMetGlnLeuGlnGlyGlnAsnGluValGluGlnArgGlnArgLysValValIleLeu 62
Db 1219385 CTCGACGCGCTGTGCTGCTCGCTGATCAACACCCCG-----GTAGATCTGCTG 1219335
Qy 63 SerGlnAspArgPheTyLysValLeuThrAlaGluGlnLysAlaLysAlaLeuLysGly 82
Db 1219334 ACCACGACGCGCTTCTC---TACCCCAACGCGGACTGACGCGGCAAACTCATGATC 1219278
Qy 83 GlnTyLysAsnPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThrLeuLys 102
Db 1219277 CCGAAAGGTTTT-----CCGAAAGCTATTAACGCGCGGCACTGATGCGTTTGACC 1219224
Qy 103 AsnIleValGluGlyLysThrValGlu---ValProThrTyLysAspPheValThrHisSer 121
Db 1219223 TCGGTAAAGTCCGGCTCCGACTACGATGCGCGCGCTGATTCGATTCGACTACGAC 1219164
Qy 122 ArgLeuPro-----GluThrThrValValTyProAlaAspValValLeuPheGluGly 139
Db 1219163 ATCATCCCGGGGACAGAACAGTGTCCGCAATCCC---GACATCTGATCTCGAGGGG 1219107
Qy 140 IleLeuValPheTySerGlnGlu-----IleArgAspMetPheHisLeuArgLeu 156
Db 1219106 CTCAAAGTCTTGGACGCGCGCCACGCTCATGCTGTGCAATCTGTTTCATTTTTCGTTG 1219047
Qy 157 PheValAspThrAsp---SerAspValArg-----LeuSerArgArgValLeuArg 172
Db 1219046 TATGTGAGCGCCCGGATCGAGATCGAGCTGACGTGATGATCAAGCTTTTGCCATG 1218987
Qy 173 AspValArgArgGlyArgAspLeuGlnIleLeuThrGlnTyLysThrThrPheVal--- 191
Db 1218986 CGCACACGCGGCTGTGCGCGACCGCGGAATCACTTCACACATACGCGGCTTTTCTGAC 1218927
Qy 192 -----LysProAlaPheGlu 196
Db 1218926 TCGCAGCGCGCTGTCGCGCGCGGAGATGTGCGGACCATCAACCGGCCCACTCGTG 1218867
Qy 197 GluPheCysLeuProThrLysLysTyLysAlaAspValIleIleProArgGlyValAspAsn 216
Db 1218866 GAGAAATCTTGTCCACCGCTGCCCGGACACCTGTGCTGTGCGCAAGACCGCATCAT 1218807
Qy 217 MetVal----- 218
Db 1218806 TCCATTAAACCGGCTGCGGCTGCGCAAGCTGTAATTGGCTCAATCGCGGGTGACAGCG 1218747

```

Qy 219 -----AlaIleasneuleuleValGlnHisIleGlnAspIleLeuAsn 232
Db 1218746 GAGTAGGCGCCGACCGCGGACCTCCGGATATTGCGATCCCTCGAAAC-----CGT 1218693
Qy 233 GlyAspIleCysLysTrpHisArgGlyLysSerAsnGlyArgSerTrpLysArgThrPhe 252
Db 1218692 GGTGTACCCCGCTCGGCGCGGCTAGCGGATACCGGTTCGGGTGATCGCGCCAAATTC 1218633
Qy 253 SerGluProGlyAspHisProGlyMet 261
Db 1218632 AGCGGCAACGATCGACACAGGCGCTG 1218606
RESULT 12
US-09-103-840A-1/c
Sequence 1, Application US/09103840A
Patent No. 6294328
GENERAL INFORMATION:
APPLICANT: FLEISCHMAN, Robert D.
APPLICANT: WHITE, Owen R.
APPLICANT: FRASER, Claire M.
APPLICANT: VENTNER, John C.
TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
TUBERCULOSIS
FILE REFERENCE: 24366-20007.00
CURRENT APPLICATION NUMBER: US/09/103.840A
CURRENT FILING DATE: 1998-06-24
NUMBER OF SEQ ID NOS: 2
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 4411529
TYPE: DNA
ORGANISM: Mycobacterium tuberculosis
OTHER INFORMATION: H37Rv
US-09-103-840A-1
Alignment Scores:
Pred. No.: 3,91e-05 Length: 4411529
Score: 156.50 Matches: 70
Percent Similarity: 39.16% Conservatave: 51
Best Local Similarity: 22.65% Mismatches: 127
Query Match: 10.79% Indels: 61
Gaps: 13
US-09-896-522-2 (1-277) x US-09-103-840A-1 (1-4411529)
Qy 3 SerIleGlyGlyIleAspCysGluSerProIleAlaAspArgProHisGlnArg 22
Db 1219971 GCGGCCACCGCGGAATTTCTCGCGAGCCCAACAAACCGGACCGCG-----GTG 1219918
Qy 23 ProPheLeuIleGlyValSerGlyGlyThrAlaSerGlyLysSerThrValCysGlnLys 42
Db 1219917 CCGTTTCATCGGTGTGTGCGCGGAGTGTGGCGGTGCGCAAGTCACACCGCGCGGTG 1219858
Qy 43 IleMetGluLeuLeuGlyIleAsnGlnValGlnGlnArgGlnArgLysValIleLeu 62
Db 1219857 CTCGAGGCGCTGCTGCGTCTGCGGATCACACCGCGG-----GTAGATCTGTG 1219807
Qy 63 SerGlnAspArgPheTrpLysValLeuThrAlaGlnGlnLysAlaLysAlaLeuLysGly 82
Db 1219806 ACCACCGACGCGCTTCTC---TACCCCAACCGCGAGCTGACGCGGCAACCTCATGCAT 1219750
Qy 83 GlnTrpAsnPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThrLeuLys 102
Db 1219749 CGCAAAAGTTT-----CCGAAAGCTATACCGCGCGGACATGATCGGTTTGTAC 1219696
Qy 103 AsnIleValGlnGlyLysThrValGln---ValProThrTrpAspPheValThrHisSer 121
Db 1219695 TCGGTGAAGTCGCGTCCGACTACGATGTCGCGGCTGTATTCGACTTTCGACTACGAC 1219636
Qy 122 ArgLeuPro-----GluThrThrValValTrpAlaAspValValLeuPheGlnGly 139
Db 1219635 ATCATCCGGGGGAGCAACAGTGTCTCGCATCC---GACATCTGATCTCTGAGGGG 1219579

Qy 140 IleLeuValPheTrpSerGlnGlu-----IleArgAspMetPheHisLeuArgLeu 156
Db 1219578 CTCAGCTCTTGAGAGACCGGCGCCAGCTCATGTGTGTGATCTTTCATTTTGTG 1219519
Qy 157 PheValAspThrAsp---SerAspValArg-----LeuSerArgArgValLeuArg 172
Db 1219518 TATGTAGACCGCCGATCGAGCATCGAGACAGTGTAGTACGATACAGGTTTTCGCCAAG 1219459
Qy 173 AspValArgArgGlyArgAspLeuGlnIleLeuThrGlnThrThrPheVal--- 191
Db 1219458 CGCACACGCGCTTCGCCCGACCGGATACACTTCCACCACTACCGGCTTTTCTGAC 1219399
Qy 192 -----LysProAlaPheGlu 196
Db 1219398 TCGCAGCGCTGCTGCGCGCGCGGAGATGTGCGGACCATCAACCGGCCAACCTCTG 1219339
Qy 197 GluPheCysLeuProThrLysLysTrpAlaAspValIleIleProArgLysValAspAsn 216
Db 1219338 GAGAACATTCTGCCGACCGCTCCCGGCGCACCTGTGCTGCGCAAGACGCGCATCAT 1219279
Qy 217 MetVal----- 218
Db 1219278 TCCATTACCGGCTGCGGCTGCGCAAGCTGTAAATTGGCGGTCCATTCGCCGCTGACGC 1219219
Qy 219 -----AlaIleAsnLeuIleValGlnHisIleGlnAspIleLeuAsn 232
Db 1219218 GAGTTAGCGCGGACCGCGGCGACCTCCGATTTTCATCTCCGAAAC-----CGT 1219165
Qy 233 GlyAspIleCysLysTrpHisArgGlyLysSerAsnGlyArgSerTrpLysArgThrPhe 252
Db 1219164 GGTGTACCCGCTCGGCGCGGCTAGCGGAGATACCGGTTCGGGTGATCGCGCCAAATTC 1219105
Qy 253 SerGluProGlyAspHisProGlyMet 261
Db 1219104 AGCGGCAACGATCGACACAGGCGCTG 1219078
RESULT 13
US-09-557-884-1/c
Sequence 1, Application US/09557884
Patent No. 6506581
GENERAL INFORMATION:
APPLICANT: Fleischmann et al.
TITLE OF INVENTION: The Nucleotide sequence of the Haemophilus influenzae Rd Genome, Fragments thereof, and Uses Thereof
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: MD
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: 3 1/2 inch diskette
COMPUTER: Dell Pentium
OPERATING SYSTEM: MS DOS v6.22
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/557,884
FILING DATE: 25-Apr-2000
CLASSIFICATION: <Unknown>
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 08/476,102
FILING DATE: JUN-5-1995
ATTORNEY/AGENT INFORMATION:
NAME: Michelle S. Marks
REGISTRATION NUMBER: 41,971
REFERENCE/DOCKET NUMBER: PB186P3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 301-309-8504
TELEFAX: 301-309-8439
INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:
 LENGTH: 1830121 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 SEQUENCE DESCRIPTION: SEQ ID NO: 1:
 US-09-557-884-1

Alignment Scores:
 Pred. No.: 0.00341 Length: 1830121
 Score: 137.50 Matches: 52
 Percent Similarity: 41.70% Conservative: 46
 Best Local Similarity: 22.13% Mismatches: 86
 Query Match: 9.48% Indels: 51
 DB: 4 Gaps: 7

US-09-896-522-2 (1-277) x US-09-557-884-1 (1-1830121)
 QY 23 ProPheLeuIleGlyValSerGlyGlyThrAlaSerGlyLysSerThrValCysGlyLys 42
 DB 670207 CCTTACATTAATTAATGCTGCGACGTGCTGTTGGGAAAAGCACTTTCGACGATTT 670148
 QY 43 IleMeGluLeuGlyGlnAsnGluValGluGlnArgGlnArgValAlaLeu 62
 DB 670147 CTACATCTCTACTTCTCATTTGCCAATCGA-----AGAAAAGTTGATCTCATC 670097
 QY 63 SerGlnAspArgPheTyrLysValLeuThrAlaGluGlnLysAlaLeuLysGly 82
 DB 670096 ACAACATGAGGTTCTTATCCATTAAACAAGCTTAAACAAGTAATCTTTTAAACAAG 670037
 QY 83 GlnTyrAsnPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThrLeuLys 102
 DB 670036 AAA-----GGTTTTCCCGTTTCTTATGACACACCTTAACCTCATTCGCTTTTACCG 669986
 QY 103 AsnIleValGluGlyLysThr--ValGluValProThrTyrAspPheValThrHisSer 121
 DB 669985 GATGTAAATCAAGGTAAAGCAATGTTACAGCACCAATTACTCTCATTTAAGAGGCTTG 669926
 QY 122 ArgLeuProGluThr--ThrValValTyrProAlaAspValAlaLeuPheGluGlyLys 140
 DB 669925 ATTAATTCCTGATTAATTTGATGAGTAAACCTCAATCTTATTTAAGAGGCTTG 669866
 QY 141 LeuValPheTyrSerGlnGluIleArg--AspMetPheHisLeuArgLeuPheValAsp 159
 DB 669865 AATGTACTTCAACGGGTATTAATAAACCGATCAACATTCGTATCAGATTTTGCGAT 669806
 QY 160 ThrAspSerAspValArgLeuSerArgValLeuArgAspValArgArgGlyArgAsp 179
 DB 669805 TTCTCTATTATGTTGATGCGAAGAAAATTATTTGAAA----- 669767
 QY 180 LeuGluGlnIleLeuThrGlnTyrThrPheValLysProAlaPhe----- 195
 DB 669766 ---GAATGTATATCAACAGATTTCTGAAATTCGCGAAGTGATTCACGATCTTAAT 669710
 QY 195 ----- 195
 DB 669709 TCTTACTTTAAACATTACGACAGCTTTCAAAAGAAAGGCTATCGGACGCAAGTAAA 669650
 QY 196 -----GluGluPheCysLeuProThrLysTyr 205
 DB 669649 ATCTGGAGTAATAATTAACGATTAATCTCAATCAAAAACATTTCTTCCAAACCGGACGCA 669590
 QY 206 AlaAspValIleLeuProArgGlyValAspAsnMetValAlaIle 220
 DB 669589 GCAAATTAAATTTAAAAAAGGTCATATACCAAGTAAAGTAAATTA 669545
 RESULT 14
 US-09-643-990A-1/c
 ; Sequence 1, Application US/09643990A
 ; Patent No. 6528289
 ; GENERAL INFORMATION:
 ; APPLICANT: Robert D. Fleischmann
 ; Mark D. Adams

Owen White
 Hamilton O. Smith
 J. Craig Venter
 TITLE OF INVENTION: The Nucleotide sequence of
 the Haemophilus influenzae Rd Genome, Fragments
 Thereof, and Uses Thereof
 NUMBER OF SEQUENCES: 1
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Human Genome Sciences, Inc.
 STREET: 9410 Key West Avenue
 CITY: Rockville,
 STATE: MD
 COUNTRY: USA
 ZIP: 20850
 COMPUTER READABLE FORM:
 MEDIUM TYPE: 3 1/2 inch diskette
 COMPUTER: Dell Pentium
 OPERATING SYSTEM: MS DOS V6.22
 SOFTWARE: ASCII Text
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/643,990A
 FILING DATE: 23-Aug-2000
 CLASSIFICATION: <Unknown>
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/487,429
 FILING DATE: 1995-06-07
 APPLICATION NUMBER: 08/426,787
 FILING DATE: 1995-04-21
 ATTORNEY/AGENT INFORMATION:
 NAME: Kenley K. Hoover
 REGISTRATION NUMBER: 40,302
 REFERENCE/DOCKET NUMBER: PB186PIC1
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 301-610-5790
 TELEFAX: 310-309-8439
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1830121 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 SEQUENCE DESCRIPTION: SEQ ID NO: 1:
 US-09-643-990A-1
 Alignment Scores:
 Pred. No.: 0.00341 Length: 1830121
 Score: 137.50 Matches: 52
 Percent Similarity: 41.70% Conservative: 46
 Best Local Similarity: 22.13% Mismatches: 86
 Query Match: 9.48% Indels: 51
 DB: 4 Gaps: 7
 US-09-896-522-2 (1-277) x US-09-643-990A-1 (1-1830121)
 QY 23 ProPheLeuIleGlyValSerGlyGlyThrAlaSerGlyLysSerThrValCysGlyLys 42
 DB 670207 CCTTACATTAATTAATGCTGCGACGTGCTGTTGGGAAAAGCACTTTCGACGATTT 670148
 QY 43 IleMeGluLeuGlyGlnAsnGluValGluGlnArgGlnArgValAlaLeu 62
 DB 670147 CTACATCTCTACTTCTCATTTGCCAATCGA-----AGAAAAGTTGATCTCATC 670097
 QY 63 SerGlnAspArgPheTyrLysValLeuThrAlaGluGlnLysAlaLeuLysGly 82
 DB 670096 ACAACATGAGGTTCTTATCCATTAAACAAGCTTAAACAAGTAATCTTTTAAACAAG 670037
 QY 83 GlnTyrAsnPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThrLeuLys 102
 DB 670036 AAA-----GGTTTTCCCGTTTCTTATGACACACCTTAACCTCATTCGCTTTTACCG 669986
 QY 103 AsnIleValGluGlyLysThr--ValGluValProThrTyrAspPheValThrHisSer 121
 DB 669985 GATGTAAATCAAGGTAAAGCAATGTTACAGCACCAATTACTCTCATTTAAGAGGCTTG 669926

```

Oy      122  rglseuProGlnThr---ThrValValItyrProAlaAspValValLeuPheGluGlyIle 140
Db      669925 ATTATTCGCCGATTAATTTTGTATGTGTGAAGATAAACCGCATCTTCTTATTTTAGAAGCCTTG 669866
Oy      141  LeuValPheTySerGInGluIleArg---AspMetPheHisLeuArgLeuPheValAsp 159
Db      669865 AATGTAACCTCAACGGGTATATATAATAAAACCGATCAAAACATTCCGTATCAGATTTTGTGAT 669808
Oy      160  ThrAspSerAspValArgLeuSerArgValLeuArgAspValArgArgGlyArgAsp 179
Db      669805 TTCTCTATTATGATTCATGACTGAGAAGAAAATTATTAATA----- 669767
Oy      180  LeuGluGlnIleLeuThrGlnItyrThrPheValIysProAlaPhe----- 195
Db      669766 ---GATGSTATATCAACAAGATTTCTGAATTCGGGAAGTCGATTCACAGATTCCTAAT 669710
Oy      195 ----- 195
Db      669709 TCTTACTTTAAACATTACGCAAGCTTATCAAAAGAGGCTATCGCAGCGCAAGTAA 669650
Oy      196 -----GluGlnPheYsIeuProHyrIysTyr 205
Db      669649 ATCTGGGATGAAATTACGAGATTAAATCTCAATCAAAAACTTTTCCAAACCGGTGAGCGA 669590
Oy      206  AlaAspValIleIleProArgGlyValAspAsnMetValAlaIle 220
Db      669589 GCAAATTTAATCTTAATAAAAGATCATATACCAAGTAGAATTA 669545

RESULT 15
US-09-221-017B-536
; Sequence 536, Application US/09221017B
; Patent No. 644799
; GENERAL INFORMATION:
; APPLICANT: ROSS, BRUCE C.
; TITLE OF INVENTION: P. GINGIVALIS NUCLEOTIDES AND USES THEREOF
; NUMBER OF SEQUENCES: 1120
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 PAGE MILL ROAD
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FASTSEQ for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/221, 017B
; FILING DATE: 23-DEC-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PP1182
; FILING DATE: 31-DEC-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PP1546
; FILING DATE: 30-JAN-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PP2911
; FILING DATE: 09-APR-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/AU98/01023
; FILING DATE: 10-DEC-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Monroy, Gladys H
; REGISTRATION NUMBER: 32,430
; REFERENCE/DOCKET NUMBER: 27340-20021.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-813-5600
; TELEFAX: 650-494-0792
; TELEX: 706141
;

```

```

? INFORMATION FOR SEQ ID NO: 536:
? SEQUENCE CHARACTERISTICS:
?     LENGTH: 973 base pairs
?     type: nucleic acid
?     STRANDEDNESS: double
?     TOPOLOGY: circular
?     MOLECULE TYPE: DNA (genomic)
?     HYPOTHEetical: NO
?     ANTI-SENSE: UNKNOWN
?     ORIGINAL SOURCE:
?     ORGANISM: PORYPHYROMONAS GINGIVALIS
?     FEATURE:
?     NAME/KEY: misc feature
?     LOCATION: 1..973
?
US-09-221-017B-536

Alignment Scores:
Pred. No.: 1.44e-06 Length: 973
Score: 122.00 Matches: 35
Percent Similarity: 46.21% Conservative: 26
Best Local Similarity: 26.52% Mismatch: 21
Query Match: 8.41% Indels: 50
DB: 4 Gaps: 4

US-09-896-522-2 (1-277) x US-09-221-017B-536 (1-973)

QY 62 LeuSerGlnAspArgPheTyrIleValLeuThrAlaGluGlnIleValAlaIleLeuLys 81
DB 27 CTTTCGCTCGAAGACATTTTC-----ATCAACCGCGAGATTCCGCGAGTAATCC 80
QY 82 GlyGluTyrAsnPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThrLeu 101
DB 81 GGAAGATTAAGATTGGAAATCGCTACGCTTGGATCTGCCCTTTTCAATAGAGATCTG 140
QY 102 LysAsnIleValGluGlyLysThrValGluValProThrTyrAspPheValThrHisSer 121
DB 141 AAGCAATGATATCGCGGTGAAGAAAGTGAAGCTGCCCACTTGTGCTTCGCAACCGGATG 200
QY 122 ArgLeu-----ProGluThrThrValValTyrProAlaAspValIleLeuPheGluGly 139
DB 122 ArgLeu-----ProGluThrThrValValTyrProAlaAspValIleLeuPheGluGly 139
QY 201 AGGGCTTCAAAAGAAATTCATACAGCTGAAGAAGCGAGTATTCCTGATCTCGAAGGC 260
DB 201 AGGGCTTCAAAAGAAATTCATACAGCTGAAGAAGCGAGTATTCCTGATCTCGAAGGC 260
QY 140 IleLeuValPheTyrSerGln-----GluIleArgAspMetPheHisLeuArgLeu 156
DB 140 IleLeuValPheTyrSerGln-----GluIleArgAspMetPheHisLeuArgLeu 156
QY 261 ATTCAAGCCCTGAATCCGAGCTGATACCCGGTGTGCCCGCAATCAAGTATATTAAGAT 320
DB 261 ATTCAAGCCCTGAATCCGAGCTGATACCCGGTGTGCCCGCAATCAAGTATATTAAGAT 320
QY 157 PheValAspThr-----AspSerAsp 163
DB 157 PheValAspThr-----AspSerAsp 163
QY 321 TATGTAGAGCGCTCTCACTGCAATAGATTGATGCTCACAATCCGATCCGACATACGAC 380
DB 321 TATGTAGAGCGCTCTCACTGCAATAGATTGATGCTCACAATCCGATCCGACATACGAC 380
QY 164 ValArgLeuSerArgArgValLeuArgAspValArg 175
DB 164 ValArgLeuSerArgArgValLeuArgAspValArg 175
DB 381 AACCGACTATCAAGGAGATGTGGTGGCGACATATATCC 416

```

Search completed: November 25, 2003, 12:12:34
Job time : 7528 secs


```
QY 121 SerArgLeuProGluThrThrValValTyrProAlaAspValValLeuPheGluGlyIle 140
Db 361 TCAGAGTTTACAGACGACCGTGTCTACCTCGGACGGTGTCTGTTGAGGGCATC 420
QY 141 LeuValPheTyrSerGlnGluIleArgAspMetPheHisLeuArgLeuPheValAspThr 160
Db 421 TTGGTGTCTTCAACCCAGAGATCCGGGACATGTTCCACCTCGCCCTCTTCTGTGACACC 480
QY 161 AspSerAspValArgLeuSerArgArgValLeuArgAspValArgArgGlyArgAspLeu 180
Db 481 GACTCCGACGTCAGGCTGTCTCGAAGAGTTCTCCGGGACGTCGCGGAGGAGGACCTG 540
QY 181 GlnGlnIleLeuThrGlnTyrThrThrPheValIysProAlaPheGluGluPheCysLeu 200
Db 541 GAGGAGATTCGACGACGATACCACTTCTGTGAAGCCGGCTTCGAGAGGATTCGCTG 600
QY 201 ProThrIysLysTyrAlaAspValIleIleProArgGlyValAspAsnMetValAlaIle 220
Db 601 CCGCAAAAGAGTATGCCGATGTATCATCCACGAGAGTGACAAATATGTTGCCATC 660
QY 221 AsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysLysTyrPheArg 240
Db 661 AACCTGATTCGTCAGCATCATCCAGACATTCGTAATGCTGACATCTCAATGACACCGA 720
QY 241 GlnGlySerAsnGlyArgSerTyrLysArgThrPheSerGluProGlyAspPheArgGly 260
Db 721 GGAAGGTCACATGGCGGAGTACAGAGGACCTTTTCTGAGCGGAGACACCCCTGGG 780
QY 261 MetLeuThrSerGlyLysArgSerHisLeuGluSerSerArgProHis 277
Db 781 ATGTGACCTCTGGCAAAAGGTCACTTTGGAGTCCAGACAGACCCCAAC 831

RESULT 2
US-10-037-270-546
; Sequence 546, Application US/10037270
; Publication No. US20030104529A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Weinman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunding
; APPLICANT: Wang, Duntui
; APPLICANT: Wang, Zhimei
; APPLICANT: Tillinghaest, John
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20030104529A1el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/10/037, 270
; CURRENT FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: pc FL_genes Version 1.0
; SEQ ID NO 546
; LENGTH: 1288
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (95)..(928)
```

```
US-10-037-270-546
Alignment Scores:
Pred. No.: 7,99e-287 Length: 1288
Score: 277.00 Matches: 277
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0

US-09-896-522-2 (1-277) x US-10-037-270-546 (1-1288)
QY 1 MetAlaSerAlaGlyGluAspCysGluSerProAlaProGluAlaAspArgProHis 20
Db 95 ATGGCTTCGGCGGAGCGCAAGCTCCGAGAGCCCGCGCGGAGCCGACGTCGCCAC 154
QY 21 GlnArgProPheLeuIleGlyValSerGlyGlyThrAlaSerGlyLysSerThrValCys 40
Db 155 CAGCGCCCTTCTCTGATAGGGGTGAGCGCGGCGGACCTGCAGCGGAAATCGACCTGTGT 214
QY 41 GlnLysIleMetCyluLeuLeuGlyGlnAsnGluValGlnGlnPheGlnArgLysValVal 60
Db 215 GAGAGATATCATGAGTTGCTGGGACAGAAAGAGTGAACAGCGGACGGAAGGTGTCT 274
QY 61 IleLeuSerGlnAspArgPheTyrLysValLeuThrAlaGlnGlnLysAlaLysAlaLeu 80
Db 275 ATCTTAGCCAGGACAGGTTCTTCAAGGTCCTGACCGGACGAGACGAAGGCGCAAGCTTG 334
QY 81 LysGlyGlnTyrAsnPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThr 100
Db 335 AAAGGACAGTACATTTTGACCATCCAGATGCCCTTGATATGATTTGATGACAGAGCT 394
QY 101 LeuLysAsnIleValGlnGlyLysThrValGlnValProThrTyrAspPheValThrHis 120
Db 395 CTAAAGATCATCTGAGAGGCGCAAAAGGTGAGAGTCCGACCTATATTTGTGACACAC 454
QY 121 SerArgLeuProGluThrThrValValTyrProAlaAspValValLeuPheGluGlyIle 140
Db 455 TCAGAGTTTACAGACGACCGTGTCTACCCCTCGCGGACGTGTTCTGTTTGAAGGCATC 514
QY 141 LeuValPheTyrSerGlnGluIleArgAspMetPheHisLeuArgLeuPheValAspThr 160
Db 515 TTGGTGTCTTCAACGACGAGAGATCCGGGACATGTTCCACTGCGCTCTTCGTGGACACC 574
QY 161 AspSerAspValArgLeuSerArgArgValLeuArgAspValArgArgGlyArgAspLeu 180
Db 575 GACTCCGACGTCAGGCTGTCTCGAAGAGTTCTCCGGGACGTGCGGAGGAGGACCTG 634
QY 181 GlnGlnIleLeuThrGlnTyrThrThrPheValIysProAlaPheGluGluPheCysLeu 200
Db 635 GAGCAGATTCGACGACGATACCACTTCTGTGAACCCGCGCTTCGAGGAGTTCTGCCCTG 694
QY 201 ProThrIysLysTyrAlaAspValIleIleProArgGlyValAspAsnMetValAlaIle 220
Db 695 CCGACAAAGAGTATCCGATGTGATCATCCCGAGAGTGAACAATGTGTTGCACTC 754
QY 221 AsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysLysTyrPheArg 240
Db 755 AACCTGATTCGTCAGCATCATCCAGACATTCGTAATGAGACATTCGCAATATGACCCGA 814
QY 241 GlnGlySerAsnGlyArgSerTyrLysArgThrPheSerGluProGlyAspHisProGly 260
Db 815 GAGAGGTCATGAGGCGGAGCTCAAGCGGACCTTTTCTGAGCAGAGGAGCACCCTGGG 874
QY 261 MetLeuThrSerGlyLysArgSerHisLeuGluSerSerArgProHis 277
Db 875 ATGTGACCTCTGGCAAAAGGTCACTTTGGAGTCCAGACAGACCCCAAC 925

RESULT 3
US-09-896-522-1
; Sequence 1, Application US/09896522
; Patent No. US2002005161A1
; GENERAL INFORMATION:
```


QY 141 LeuValPheTyrSerGlnGluIleArgAspMetPheHisLeuArgLeuPheValAspThr 160
DB 504 TTGGTGTCTTACACCGACGAGATCCGGACATGTTCCACCTTCCTTCGTGGACACC 563
QY 161 AspSerAspValaLysSerArgArGValLeuArgAspValaArgAGGlyATArgAspLeu 180
DB 564 GACTCCCAAGTCACGCTGTCTCGAAGAGTTCTCCGGAGCTGGCCGAGGAGGAGGAGCTTG 623
QY 181 GlnGlnIleLeuThrGlnTyrThrThrPheValaLysProAlaPheGlnGluPheCysLeu 200
DB 624 GAGCAGATTCTGACGACATCACACCTTCGTGAAGCGGCTTCGAGGAGTTCTGCTG 683
QY 201 ProThrTyrLysTyr-AlaAspValaLleProArgGlyValaAspMetValaAla 220
DB 684 CCGCAAGAAGATGTTCCGATGATGATCCCGAGAGATGACAAATGTGTTGCCAT 743
QY 220 eAsnLeuIleValaGlnHisIleGlnAspIleLeuAsnGlyAspIleCysLysTrpHis 240
DB 744 CAACCTGATCTGACGACATCCGACATTCGAAATGGTGACATCTGCAAAATGGCACCG 803
QY 240 GGLyGlySerAsnGlyArgSerTyrLysArgThrPheSerGluProGlyAspHisProGly 260
DB 804 AGGAGGGTCCATGGCGGAGACTCAAGCGGACCTTTTCTGAGCCAGGAGGACCACTGG 863
QY 260 yMetLeuThrSerGlyLysArgSerHisLeuGluSerSerSerArgProHis 277
DB 864 GATGCTGACCTTGGCAAAACGGTCACATTTGGAGTCCAGCAGACACCCAC 915

RESULT 5

US-09-918-995-30379
Sequence 30379, Application US/09918995
Publication No. US20030073623A1
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc.
TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
FILE REFERENCE: 20411-756
CURRENT APPLICATION NUMBER: US/09/918, 995
CURRENT FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: US/09/235, 076
PRIOR FILING DATE: 1999-01-20
NUMBER OF SEQ ID NOS: 38054
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 30379
LENGTH: 472
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)..(472)
OTHER INFORMATION: n = A,T,C or G
US-09-918-995-30379

Alignment Scores:

Pred. No.: 5,16e-134 Length: 472
Score: 134.00 Matches: 134
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 48.38% Indels: 0
DB: 11 Gaps: 0

US-09-896-522-2 (1-277) x US-09-918-995-30379 (1-472)

QY 33 AlaSerGlyLysSerThrValaCysGluLysIleMetGluLeuGluGlnAsnGluVal 52
DB 41 GCCAGCGGGAAGTCGACCGTGTGTGAAGAATCATGGAGTTGCTGGGACGAAGAAGTG 100
QY 53 GlnGlnArgGlnaGlyLysValaValaIleLeuSerGlnAspArgPheTyrLysValaLeuThr 72
DB 101 GAACAGCGGCGAGCGAGAGTGTCATCTGAGCCAGCAGAGTTCTCAAGAGTCTGACG 160
QY 73 AlaGlnGlnLysAlaLysAlaLeuLysGlnTyrAsnPheAspHisProAspAlaPhe 92

DB 161 GCAGACGAGAAGGCCAAGGCTTTGAAAGACATGACATTTTACCATCCAGATGCTTT 220
QY 93 AspAsnAspLeuMetHisArgThrLeuLysAsnIleValaGlnGlyLysThrValaGluVal 112
DB 221 GATPAGATTTGATGACAGACTCTGAAGAATCATGCTGGAAGGCAAAACGGTGAGTG 280
QY 113 ProThrTyrAspPheValaThrHisSerArgLeuProGluThrThrValaTyrProAla 132
DB 281 CCACCTATGATTTTGTGACACACTCAAGGTTACCAAGACCAAGGATGATCCGCTGAG 340
QY 133 AspValaIleLeuPheGlnGlyIleLeuValaPheTyrSerGlnGluIleArgAspMetPhe 152
DB 341 GACTGCTTCTGTTTAAAGGACATCTTGCTTCTAAGCCAGAGATCCGGGACATGTT 400
QY 153 HisLeuArgLeuPheValaAspThrAspSerAspValaArgLeu 166
DB 401 CACTGCGCTTCTGTTGACACCGACTCCGACGTGAGGCTG 442

RESULT 6

US-10-029-386-15883
Sequence 15883, Application US/10029386
Publication No. US20030194704A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharron G.
APPLICANT: Rank, David R.
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR GI
TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
FILE REFERENCE: AEOMICA-X-2
CURRENT APPLICATION NUMBER: US/10/029,386
CURRENT FILING DATE: 2001-12-20
NUMBER OF SEQ ID NOS: 34288
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 15883
LENGTH: 187
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO CHR9.1
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.5
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.8
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.6
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.7
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.5
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.7
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.3
OTHER INFORMATION: EST HUMAN HIT: A1992171.1, EVALUE 1.00e-100
OTHER INFORMATION: SWISSPROT HIT: P52623, EVALUE 4.00e-26
OTHER INFORMATION: NT HIT: g114783235, EVALUE 1.00e-100
US-10-029-386-15883

Alignment Scores:

Pred. No.: 4,27e-55 Length: 187
Score: 60.00 Matches: 60
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 21.66% Indels: 0
DB: 12 Gaps: 0

US-09-896-522-2 (1-277) x US-10-029-386-15883 (1-187)

QY 218 ValaIleAsnLeuIleValaGlnHisIleGlnAspIleLeuAsnGlyAspIleCysLys 237
DB 3 GTTGCCATCAACTGATCGGACGACATCCAGAGATCTGAAATGATGATGATCTCAAA 62
QY 238 ThrHisArgGlyLysSerAsnGlyArgSerTyrLysArgThrPheSerGluProGlyAsp 257
DB 63 TGGCACCAGAGAGAGGTCCTCAATGGCGGAGCTCAACACCGGACCTTTCTGAGCCAGGAGAC 122
QY 258 HisProGlyMetLeuThrSerGlyLysArgSerHisLeuGluSerSerSerArgProHis 277

Db 123 CACCTGGATGCTGACCTCTGGCAACGGTCACTTTGGATGTCAGACAGACCCAC 182

RESULT 7

US-10-029-386-2183

Sequence 2183, Application US/10029386

Publication No. US20030194704A1

GENERAL INFORMATION:

APPLICANT: Penn, Sharon G.

APPLICANT: Rank, David R.

APPLICANT: Hanzel, David K.

TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G

FILE REFERENCE: AEOMICA-X-2

CURRENT APPLICATION NUMBER: US/10/029,386

CURRENT FILING DATE: 2001-12-20

NUMBER OF SEQ ID NOS: 34288

SOFTWARE: Annomax Sequence Listing Engine vers. 1.1

SEQ ID NO 2183

LENGTH: 510

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

OTHER INFORMATION: MAP TO CHR9.1

OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.5

OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.8

OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.6

OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.7

OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3

OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.5

OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.7

OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.3

OTHER INFORMATION: NT HIT: BFC64526.1, EVALUE 0.00e+00

OTHER INFORMATION: EST HUMAN HIT: BFC64526.1, EVALUE 0.00e+00

OTHER INFORMATION: SWISSPROT HIT: P52623, EVALUE 2.00e-25

US-10-029-386-2183

Alignment Scores:

Pred. No.:	1,05e-54	Length:	510
Score:	60.00	Matches:	60
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	21.66%	Indels:	0
DB:	12	Gaps:	0

US-09-896-522-2 (1-277) x US-10-029-386-2183 (1-510)

Qy 218 ValAlaIleAsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysLys 237

Db 26 GTTGCCATCAACCTGATCGTGAGCAGACATCCAGACATTTGATGATGTCATCTGCATA 85

Qy 238 TrpHisArgGlyGlySerAsnGlyArgSerTyrIleArgThrPheSerGluProGlyAsp 257

Db 86 TGGCACCAGAGAGGCTCCATGCGGAGCTACAGAGGAGCCTTTCTTGAGCCAGGGGAC 145

Qy 258 HisProGlyMetLeuThrSerGlyIleArgSerHisLeuGlnSerSerArgProHis 277

Db 146 CACCTGGAGTGTGACCTCTGGCAACGGTCACTTTGGATGTCAGACAGACCCAC 205

RESULT 8

US-09-764-877-2804

Sequence 2804, Application US/09764877

Patent No. US20020147140A1

GENERAL INFORMATION:

APPLICANT: Rosen et al.

TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

FILE REFERENCE: PC005

CURRENT APPLICATION NUMBER: US/09/764,877

CURRENT FILING DATE: 2001-01-17

Prior application data removed - refer to PALM or file wrapper

NUMBER OF SEQ ID NOS: 4031

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 2804

LENGTH: 9732

TYPE: DNA

ORGANISM: Homo sapiens

US-09-764-877-2804

Alignment Scores:

Pred. No.:	1.49e-53	Length:	9732
Score:	60.00	Matches:	60
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	21.66%	Indels:	0
DB:	10	Gaps:	0

US-09-896-522-2 (1-277) x US-09-764-877-2804 (1-9732)

Qy 218 ValAlaIleAsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysLys 237

Db 3901 GTTGCCATCAACCTGATCGTGAGCAGACATCCAGACATTTGATGATGTCATCTGCATA 3960

Qy 238 TrpHisArgGlyGlySerAsnGlyArgSerTyrIleArgThrPheSerGluProGlyAsp 257

Db 3961 TGGCACCAGAGAGGCTCCATGCGGAGCTACAGAGGAGCCTTTCTTGAGCCAGGGGAC 4020

Qy 258 HisProGlyMetLeuThrSerGlyIleArgSerHisLeuGlnSerSerArgProHis 277

Db 4021 CACCTGGAGTGTGACCTCTGGCAACGGTCACTTTGGATGTCAGACAGACCCAC 4080

RESULT 9

US-09-764-877-2805

Sequence 2805, Application US/09764877

Patent No. US20020147140A1

GENERAL INFORMATION:

APPLICANT: Rosen et al.

TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

FILE REFERENCE: PC005

CURRENT APPLICATION NUMBER: US/09/764,877

CURRENT FILING DATE: 2001-01-17

Prior application data removed - refer to PALM or file wrapper

NUMBER OF SEQ ID NOS: 4031

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 2805

LENGTH: 19125

TYPE: DNA

ORGANISM: Homo sapiens

US-09-764-877-2805

Alignment Scores:

Pred. No.:	2.75e-53	Length:	19125
Score:	60.00	Matches:	60
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	21.66%	Indels:	0
DB:	10	Gaps:	0

US-09-896-522-2 (1-277) x US-09-764-877-2805 (1-19125)

Qy 218 ValAlaIleAsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysLys 237

Db 5997 GTTGCCATCAACCTGATCGTGAGCAGACATCCAGACATTTGATGATGTCATCTGCATA 6056

Qy 238 TrpHisArgGlyGlySerAsnGlyArgSerTyrIleArgThrPheSerGluProGlyAsp 257

Db 6057 TGGCACCAGAGAGGCTCCATGCGGAGCTACAGAGGAGCCTTTCTTGAGCCAGGGGAC 6116

Qy 258 HisProGlyMetLeuThrSerGlyIleArgSerHisLeuGlnSerSerArgProHis 277

Db 6117 CACCTGGAGTGTGACCTCTGGCAACGGTCACTTTGGATGTCAGACAGACCCAC 6176

RESULT 10

US-09-925-300-220

Sequence 220, Application US/09925300

Patent No. US20020151681A1

GENERAL INFORMATION:

APPLICANT: Craig Rosen,

```

; APPLICANT: Steve Ruben
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05968
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 220
; LENGTH: 1310
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-925-300-220

Alignment Scores:
Pred. No.: 7,73e-17 Length: 1310
Score: 25.00 Matches: 25
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 9.03% Indels: 0
DB: 10 Gaps: 0

US-09-896-522-2 (1-277) x US-09-925-300-220 (1-1310)

Qy 189 ThrPheVallyspProlaPheGluGluPheCysLeuProthrlYsTYrAlaAspVal 208
Db 789 ACGTTCGTCMAAGCCTGCTTGAGGAATTCGTTGCCAACAAAGATATGCTGATGTG 848

Qy 209 IleIleProArgGly 213
Db 849 ATCATCCCTAGAGGT 863

RESULT 11
US-10-098-841-53
; Sequence 53, Application US/10098841
; Publication No. US20020197679A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Xu, Chongjun
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yungqing
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhao, Qing A.
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Wang, Duntui
; APPLICANT: Wang, Zhiwei
; APPLICANT: Wehrman, Tom
; APPLICANT: Zhang, Jie
; APPLICANT: Qian, Xiaohong B.
; APPLICANT: Dimaac, Radoje T.
; TITLE OF INVENTION: No. US20020197679A1el Nucleic Acids and
; FILE REFERENCE: 784CIP2
; CURRENT APPLICATION NUMBER: US/10/098,841
; CURRENT FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: 09/598,042
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 331
; SOFTWARE: pc_Fl_genes Version 1.0
; SEQ ID NO 53
; LENGTH: 1402
; TYPE: DNA
; ORGANISM: Homo sapiens
```

```

; FEATURE:
; NAME/KEY: CDS
; LOCATION: (230) ..(1075)
US-10-098-841-53

Alignment Scores:
Pred. No.: 8.21e-17 Length: 1402
Score: 25.00 Matches: 25
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 9.03% Indels: 0
DB: 13 Gaps: 0

US-09-896-522-2 (1-277) x US-10-098-841-53 (1-1402)

Qy 189 ThrPheVallyspProlaPheGluGluPheCysLeuProthrlYsTYrAlaAspVal 208
Db 848 ACGTTCGTCMAAGCCTGCTTGAGGAATTCGTTGCCAACAAAGATATGCTGATGTG 907

Qy 209 IleIleProArgGly 213
Db 908 ATCATCCCTAGAGGT 922

RESULT 12
US-09-908-975-398
; Sequence 398, Application US/09908975
; Publication No. US20030165843A1
; GENERAL INFORMATION:
; APPLICANT: SHOSHAN, Avi
; APPLICANT: WASSERMAN, Alon
; APPLICANT: MINTZ, Eli
; APPLICANT: MINTZ, Liat
; APPLICANT: FAIGLER, Simchon
; TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICE
; FILE REFERENCE: 36688-0005
; CURRENT APPLICATION NUMBER: US/09/908,975
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: US 60/287,724
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: US 60/221,607
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 32337
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 398
; LENGTH: 65
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; US-09-908-975-398

Alignment Scores:
Pred. No.: 2.68e-07 Length: 65
Score: 15.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.42% Indels: 0
DB: 12 Gaps: 0

US-09-896-522-2 (1-277) x US-09-908-975-398 (1-65)

Qy 218 ValAlaIleAsnLeuIleValGlnHisIleGlnAspIleLeuAsn 232
Db 20 GTGGCATTCACTCATCTGTCGACGATATACGGACATCTCTGAC 64

RESULT 13
US-09-908-975-24159
; Sequence 24159, Application US/09908975
; Publication No. US20030165843A1
; GENERAL INFORMATION:
; APPLICANT: SHOSHAN, Avi
; APPLICANT: WASSERMAN, Alon
; APPLICANT: MINTZ, Eli
; APPLICANT: MINTZ, Liat
```

APPLICANT: FAIGLER, Simchon
TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICE
FILE REFERENCE: 36686-0005
CURRENT APPLICATION NUMBER: US/09/908,975
CURRENT FILING DATE: 2001-07-20
PRIOR APPLICATION NUMBER: US 60/287,724
PRIOR FILING DATE: 2001-05-02
PRIOR APPLICATION NUMBER: US 60/221,607
PRIOR FILING DATE: 2000-07-28
NUMBER OF SEQ ID NOS: 32337
SOFTWARE: PatentIn version 3.0
SEQ ID NO 24159
LENGTH: 65
TYPE: DNA
ORGANISM: Mus musculus
US-09-908-975-24159

Alignment Scores:
Pred. No.: 2,686-07 Length: 65
Score: 15.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.42% Indels: 0
DB: 12 Gaps: 0

US-09-896-522-2 (1-277) x US-09-908-975-24159 (1-65)

OY 160 ThrAspSerAspValArgLeuSerArgArgValLeuArgAspVal 174
Db 1 ACAGACTCGATGTAGCGCTGTCTCGAAGAGTTCTCCGGATGTGG 45

RESULT 14
US-10-066-543-1843
Sequence 1843, Application US/10066543
Publication No. US2003087818A1
GENERAL INFORMATION:
APPLICANT: Jiang, Yuhui
APPLICANT: Pyle, Ruth A.
APPLICANT: Xu, Jiangchun
APPLICANT: Indritas, Carol Yoseph
APPLICANT: Lodes, Michael J.
APPLICANT: Secrist, Heather
APPLICANT: Carter, Derrick
APPLICANT: Fanger, Gary R.
APPLICANT: Smith, Carole L.
APPLICANT: Durham, Margarita
APPLICANT: Stolk, John A.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
FILE REFERENCE: 210121.563
CURRENT APPLICATION NUMBER: US/10/066,543
CURRENT FILING DATE: 2002-01-31
NUMBER OF SEQ ID NOS: 3417
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1843
LENGTH: 447
TYPE: DNA
ORGANISM: Homo sapiens
US-10-066-543-1843

Alignment Scores:
Pred. No.: 1,526-06 Length: 447
Score: 15.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.42% Indels: 0
DB: 14 Gaps: 0

US-09-896-522-2 (1-277) x US-10-066-543-1843 (1-447)

OY 23 ProPheLeuIleGlyValSerGlyGlyThrAlaSerGlyIysSer 37
|||||

Db 71 CCCTTCCTATAGCGCTACGCGGGGACACAGCTAGCGCAAGTCT 115

RESULT 15
US-09-918-995-23923
Sequence 23923, Application US/09918995
Publication No. US20030073623A1
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc.
TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
FILE REFERENCE: 20411-756
CURRENT APPLICATION NUMBER: US/09/918,995
CURRENT FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: US/09/235,076
PRIOR FILING DATE: 1999-01-20
NUMBER OF SEQ ID NOS: 38054
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 23923
LENGTH: 455
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(455)
OTHER INFORMATION: n = A,T,C or G
US-09-918-995-23923

Alignment Scores:
Pred. No.: 1,546-06 Length: 455
Score: 15.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.42% Indels: 0
DB: 11 Gaps: 0

US-09-896-522-2 (1-277) x US-09-918-995-23923 (1-455)

OY 23 ProPheLeuIleGlyValSerGlyGlyThrAlaSerGlyIysSer 37
Db 85 CCCTTCCTATAGCGCTACGCGGGGACACAGCTAGCGCAAGTCT 129

Search completed: November 25, 2003, 12:17:21
Job time : 379 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: November 25, 2003, 08:12:16 ; Search time 362 Seconds
(without alignments)
2521.075 Million cell updates/sec

Title: US-09-896-522-2

Perfect score: 1450
Sequence: 1 MASAGEDESPAPADRPB.....HPGMLTSGKSHLESSSRPH 277

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2190069 seqs, 1647345023 residues

Total number of hits satisfying chosen parameters: 4380138

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame+g2n.model -DEV=x1h
-Q=/cg2_1/USPTO.sp001/US09896522/rnuc.2112003.184356.5076/app.query.fasta.1.455
-DB=Published Applications NA -QFMT=fastap -SUFFIX=p2n.rmpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bases -START=1 -END=-1 -MATRIX=bloms62
-TRANS=human40.csl -LIST=45 -DOCALLIGN=200 -THR SCORE=oct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US09896522 @CGN_1.1.347 @rnuc.2112003.184356.5076
-NCPU=6 -ICPU=3 -NO MMAP -LARGESQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-Fgapop=6 -Fgapext=7 -Ygapop=10 -Ygapext=0.5 -DELop=6 -DELext=7

Database: Published Applications NA:*

1: /cg2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cg2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cg2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cg2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cg2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cg2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cg2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cg2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cg2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
10: /cg2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
11: /cg2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
12: /cg2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
13: /cg2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
14: /cg2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
15: /cg2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
16: /cg2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
17: /cg2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1450	100.0	834	9	US-09-896-522-3 Sequence 3, Appl1

2	1450	100.0	1288	14	US-10-037-270-546	Sequence 546, App
3	1450	100.0	1624	9	US-09-896-522-1	Sequence 1, Appl1
4	1437	99.1	1648	10	US-09-833-381-2048	Sequence 2048, Ap
5	940	64.8	1402	13	US-10-098-841-53	Sequence 53, Appl
6	902.5	62.2	1310	10	US-09-925-300-220	Sequence 220, App
7	727	50.1	472	11	US-09-918-995-30379	Sequence 30379, A
8	647	44.6	19125	10	US-09-764-877-2805	Sequence 2805, Ap
9	544	37.5	1401	10	US-09-938-842A-1943	Sequence 1943, Ap
10	540.5	37.3	2058	12	US-10-251-166-14	Sequence 14, Appl
11	537.5	37.1	472	11	US-09-918-995-24042	Sequence 24042, A
12	521.5	36.0	447	14	US-10-066-543-1843	Sequence 1843, Ap
13	511.5	35.3	455	11	US-09-918-995-23923	Sequence 23923, A
14	427	29.4	3662	10	US-09-070-927A-215	Sequence 215, App
15	384	26.5	1830121	14	US-10-339-960-1	Sequence 1, Appl
16	372.5	25.7	470	11	US-09-918-995-9667	Sequence 9667, App
17	351.5	24.2	15051	12	US-09-960-886-11	Sequence 11, Appl
18	351.5	24.2	15051	12	US-09-960-870-11	Sequence 11, Appl
19	351.5	24.2	580073	12	US-10-205-220-1	Sequence 1, Appl1
20	327	22.6	187	12	US-10-029-386-15883	Sequence 15883, A
21	327	22.6	510	12	US-10-029-386-2183	Sequence 2183, Ap
22	327	22.6	9732	10	US-09-764-877-2804	Sequence 2804, Ap
23	212	14.6	514	8	US-08-781-986A-955	Sequence 955, App
24	205.5	14.2	1188	10	US-09-938-842A-615	Sequence 615, App
25	199.5	13.8	926	9	US-09-770-445-403	Sequence 403, App
26	194.5	13.4	253	10	US-09-974-300-626	Sequence 626, App
27	177.5	12.2	942	10	US-09-974-300-1001	Sequence 1001, Ap
28	175.5	12.0	328	10	US-09-974-300-609	Sequence 609, App
29	174.5	12.0	263	9	US-09-923-876-5130	Sequence 5130, Ap
30	173	11.9	248	9	US-09-923-876-5906	Sequence 5906, Ap
31	154.5	10.7	936	10	US-09-738-626-1109	Sequence 1109, Ap
32	154.5	10.7	3309400	10	US-09-738-626-1	Sequence 1, Appl1
33	150.5	10.4	9025608	14	US-10-156-761-1	Sequence 1, Appl1
34	149.5	10.3	729	8	US-08-781-986A-995	Sequence 995, App
35	149	10.3	967	14	US-10-156-761-945	Sequence 945, Ap
36	137.5	9.5	1830121	14	US-10-329-960-1	Sequence 1, Appl1
37	123.5	8.4	4029	10	US-09-070-927A-298	Sequence 298, App
38	123.5	8.5	339	10	US-09-878-574-4	Sequence 4, Appl1
39	119.5	8.2	433	9	US-09-734-569-87	Sequence 87, Appl
40	117	8.1	401	10	US-09-736-457-1476	Sequence 1476, Ap
41	117	8.1	401	10	US-09-902-947-1476	Sequence 1476, Ap
42	117	8.1	401	10	US-09-849-626-1476	Sequence 1476, Ap
43	117	8.1	401	12	US-10-113-874-1476	Sequence 1476, Ap
44	117	8.1	401	14	US-10-017-754-1476	Sequence 1476, Ap
45	111.5	7.7	1425	10	US-09-938-842A-2093	Sequence 2093, Ap

ALIGNMENTS

RESULT 1
US-09-896-522-3
; Sequence 3, Application US/09896522
; Patent No. US20020055161A1
; GENERAL INFORMATION:
; APPLICANT: Glucksmann, Maria A.
; TITLE OF INVENTION: 57658, A NOVEL HUMAN URIDINE KINASE AND
; FILE REFERENCE: 381552001700
; CURRENT APPLICATION NUMBER: US/09/896, 522
; PRIOR FILING DATE: 2001-06-28
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 834
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-896-522-3

Alignment Scores:
Pred. No.: 2,81e-175
Score: 1450.00
Percent Similarity: 100.00%

Length: 834
Matches: 277
Conservative: 0

QY 161 AspSerAspValArgLeuSerArgValLeuArgAspValArgArgValArgAspLeu 180
 Db 575 GACTCCGACGTCAAGGCTGTCTCGAAGAGTTCTCCGGACGTGCGCCAGGAGGACCTG 634
 QY 181 GluGlnIleuThrGlnTyrThrThrPheValIysProAlaPheGluGluPheCysLeu 200
 Db 635 GAGCAGATTCGACGCGACGACACACCTTCGTGAAGCGGCGCTTCGAGGAGTTCTGCTG 694
 QY 201 ProThrIysIysTyrAlaAspValIleIleProArgIlyValAspAsnMetValAlaIle 220
 Db 695 CCGACAAAGAGATGTCGATGATGATCATCCACGAGAGTGAACATATGTTGCATC 754
 QY 221 AsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysIysTrpHisArg 240
 Db 755 AACCTGATCGTGCAGCATCTCGAGCATTTCTGATGTCATCTGCAATGGACCGA 814
 QY 241 GlyGlySerAsnGlyArgSerTyrIysArgThrPheSerGluProGlyAspHisProGly 260
 Db 815 GAGGGGTCCATGGCGGACGACCTACCAAGCGACCTTTCTAGCCAGGAGGACCACTGGG 874
 QY 261 MetLeuThrSerGlyIysArgSerHisLeuGluSerSerArgProHis 277
 Db 875 ATGCTGACCTCTGGCAACGATCACATTGAGTCCAGCAGACAGACCCAC 925

RESULT 3

US-09-896-522-1
 ; Sequence 1, Application US/09896522
 ; Patent No. US20020055161A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Glucksmann, Maria A.
 ; TITLE OF INVENTION: 57658, A NOVEL HUMAN URIDINE KINASE AND
 ; TITLE OF INVENTION: USES THEREOF
 ; FILE REFERENCE: 381552001700
 ; CURRENT APPLICATION NUMBER: US/09/896,522
 ; CURRENT FILING DATE: 2001-06-28
 ; PRIOR APPLICATION NUMBER: 60/216,503
 ; PRIOR FILING DATE: 2000-06-30
 ; NUMBER OF SEQ ID NOS: 6
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 1
 ; LENGTH: 1624
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (94)...(927)
 ; US-09-896-522-1

Alignment Scores:

Pred. No.: 7,76e-175 Length: 1624
 Score: 1450.00 Matches: 277
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 9 Gaps: 0

US-09-896-522-2 (1-277) x US-09-896-522-1 (1-1624)

QY 1 MetAlaSerAlaGlyGlyGluAspCysGluSerProAlaProGluAlaAspArgProHis 20
 Db 94 ATGGCTTCGGCGGAGCGGAGAGCTCGAGAGCCCGCGCGGAGCGGACCGTCCGCAC 153
 QY 21 GlnArgProPheLeuIleGlyValSerGlyIlyThrAlaSerGlyIysSerThrValCys 40
 Db 154 CAGCGGCGCTTCGTGAAGGGGTGAGCGGCGGACCTGCGCGGAGAGTGCACCGTGTGT 213
 QY 41 GluIlyIleMetCysLeuLeuGlyGlnAsnGlyValGluGlnArgGlnArgIysValVal 60
 Db 214 GAGAGATGATGAGTGTCTGGGACAGAGAGAGTGAAGAGCGGACCGGAGAGTGTGTC 273
 QY 61 IleLeuSerGlnAspArgPheTyrIysValLeuThrAlaGluGlnIlyValIysAlaLeu 80
 Db 274 ATCTTGAGCGCAGACAGTTCTCAAGGTCTGACGCGAGAGAGAGCAAGCGCTTG 333

QY 81 IysGlyGlnTyrAsnPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThr 100
 Db 334 AAGAGACGTACATTTTGAACATCCAGATGCTTTGATATATATTTGATGACAGAGACT 393
 QY 101 LeuIysAsnIleValGluGlyIysThrValGluValProThrTyrAspPheValIleHis 120
 Db 394 CTGAGAGACATGCTGAGGAGGCAAAAGCGTGAAGTCCGACCTATATATTTTGGACACAC 453
 QY 121 SerArgLeuProGluThrThrValIlyTyrProAlaAspValIleLeuPheGluGlyIle 140
 Db 454 TCAAGTTACAGAGACCAAGCGGTGTCTACCTCGCGAGAGCGTGTCTGTTTGAAGGCATC 513
 QY 141 LeuValPheTyrSerGlnGluIlyLeuArgAspMetPheHisLeuArgLeuPheValAspThr 160
 Db 514 TTGGTTTCTACAGCAGAGGATCCGGAGATGTTCCACCTGCGCTTCTGTTGAGACACC 573
 QY 161 AspSerAspValArgLeuSerArgValIleuArgAspValArgArgIysArgAspLeu 180
 Db 574 GACTCCGACGTCAAGGCTGTCTCGAAGAGTTCTCCGGACGTGCGCCAGGAGGACCTG 633
 QY 181 GluGlnIleuThrGlnTyrThrThrPheValIysProAlaPheGluGluPheCysLeu 200
 Db 634 GACGAGATTCGACGCGACGACACACCTTCGTGAAGCGGCGCTTCGAGGAGTTCCTG 693
 QY 201 ProThrIysIysTyrAlaAspValIleIleProArgIlyValAspAsnMetValAlaIle 220
 Db 694 CCGACAAAGAGATGTCGATGATGATCATCCACGAGAGTGAACATATGTTGCATC 753
 QY 221 AsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysIysTrpHisArg 240
 Db 754 AACCTGATCGTGCAGCATCTCGAGCATTTCTGATGTCATCTGCAATGGACCGA 813
 QY 241 GlyGlySerAsnGlyArgSerTyrIysArgThrPheSerGluProGlyAspHisProGly 260
 Db 814 GAGGGGTCCATGGCGGACGACCTACCAAGCGACCTTTCTAGAGCAGGAGGACCACTGGG 873
 QY 261 MetLeuThrSerGlyIysArgSerHisLeuGluSerSerArgProHis 277
 Db 874 ATGCTGACCTCTGGCAACGATCACATTGAGTCCAGCAGACAGACCCAC 924

RESULT 4

US-09-833-381-2048
 ; Sequence 2048, Application US/09833381
 ; Patent No. US20020132090A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Robison, Keith E.
 ; TITLE OF INVENTION: NO. US20020132090A1 Nucleic Acid and Protein Homologs
 ; FILE REFERENCE: 5800-119
 ; CURRENT APPLICATION NUMBER: US/09/833,381
 ; CURRENT FILING DATE: 2001-04-11
 ; PRIOR APPLICATION NUMBER: 09/516,448
 ; PRIOR FILING DATE: 2000-02-29
 ; NUMBER OF SEQ ID NOS: 2050
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 2048
 ; LENGTH: 1648
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (1)...(1648)
 ; OTHER INFORMATION: n = A, T, C or G
 ; US-09-833-381-2048

Alignment Scores:

Pred. No.: 3,64e-173 Length: 1648
 Score: 1437.00 Matches: 277
 Percent Similarity: 99.64% Conservative: 0
 Best Local Similarity: 99.64% Mismatches: 1
 Query Match: 99.10% Indels: 0
 DB: 10 Gaps: 0

US-09-896-522-2 (1-277) x US-09-833-381-2048 (1-1648)

QY 1 MetAlaSerAlaGlyGlyIuAspCyseGluSerProAlaProGluAlaAspArgProHis 20

Db 84 ATGGCTTCGGCGGGAGCGAAGACTCGAGAGCCCGCGCGAGCGGACCGTCCGAC 143

QY 21 GlnArgProHeuLeuIleGlyValSerGlyGlyThrAlaSerGlyValSerThrValCys 40

Db 144 CAGGGGCTTCCTGATAGGGGTAGCGCGGCACTCCAGCGGGAGAGTGCACCGTGT 203

QY 41 GlnUysIleMetGluLeuGlyGlnAlaSerGlyGlnAlaGlnArgGlnArgValVal 60

Db 204 GAGAGATCATGAGTGTCTGCGACAAACAGAGTGAACAGCGCGACCGGAGGTGTC 263

QY 61 IleLeuSerGlnAspArgPheTyrIleValIleuThrAlaGluGlnIleValAlaLeu 80

Db 264 ATCTGAGCCAGGACGAGCTTCACAGAGTCTTCAGCGGACAGGAGCGGACCGCTTG 323

QY 81 LysGlyGlnTyrAsnPheAspHisProAspAlaPheAspAsnAlaMetHisArgThr 100

Db 324 AAAGGACAGTACATTTTGAACCATCCAGATCCCTTGAATGATTTGATGCACAGACT 383

QY 101 LeuUysAsnIleValGlnGlyIleThrValGlnValProThrTyrAspPheValThrHis 120

Db 384 CTGAGAGACATCGTGAAGGCGAAGACGCTGAGGCTGCGACCTATGATTTTGTACACAC 443

QY 121 SerArgLeuProGluThrThrValValTyrProAlaAspValIleLeuPheGluGlyIle 140

Db 444 TCAAGGTTCACAGACACGCGGTGTACCTCGGACGCGGTCTGTGTTGAGGCGCATC 503

QY 141 LeuValPheTyrSerGlnGlnIleArgAspPhePheHisLeuArgLeuPheValAspThr 160

Db 504 TTGGTGTTCACACCGAGATCCGGACATGTTCCACCTGCGCCCTCTTGTGAGACAC 563

QY 161 AspSerAspValArgLeuSerArgValIleuArgAspValArgArgGlyArgAspLeu 180

Db 564 GACTCCGACGTCAAGCTGTCTCGAAGAGTCTCCGGGACGTCGCGGAGGAGGAGGACTG 623

QY 181 GluGlnIleLeuThrGlnTyrThrThrPheValIleAspProAlaPheGluGluPheCysLeu 200

Db 624 GAGCAGATTCGACGAGTACACACACCTTCGTGAAGCGCGCTTCGAGGAGTTCGCTG 683

QY 201 ProThrIleValSerTyrAlaAspValIleIleProArgGlyValAlaAspMetValAla 220

Db 684 CCGCAAGAGATGATGCGCATGATCATCCAGAGAGTGAACATATGTTGCCAT 743

QY 220 AsnIleuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysIleThrHisArg 240

Db 744 CAACCTGATCGTGCAGACATCCAGGACATTCGTAATGATGATCTGCAATATGCGACCG 803

QY 240 GgIyGlySerArgGlyArgSerTyrIleArgThrPheSerGluProGlyAspHisProGly 260

Db 804 AGGAGGATTCGATGAGGCGGAGCTTCAGACGCGACCTTTTTCGAGCGGAGCGACCTCG 863

QY 260 YMetLeuThrSerGlyIleAspArgSerHisLeuGluSerSerArgProHis 277

Db 864 GATCTGACTTCGCAAGCGGTCACTTTGAGTCCAGCAGACCCAC 915

RESULT 5

US-10-098-841-53

; Sequence 53, Application US/10098841

; Publication No. US20020197679A1

; GENERAL INFORMATION:

; APPLICANT: Tang, Y. Tom

; APPLICANT: Liu, Chenghua

; APPLICANT: Asundi, Vinod

; APPLICANT: Xu, Chongjun

; APPLICANT: Zhou, Ping

; APPLICANT: Ma, Yungqing

; APPLICANT: Wang, Jian-Rui

; APPLICANT: Zhao, Qing A.

; APPLICANT: Ren, Feiyan

; APPLICANT: Chen, Rui-hong

; APPLICANT: Wang, Dunrui

; APPLICANT: Wang, Zhiwei

; APPLICANT: Wehrman, Tom

; APPLICANT: Zhang, Jie

; APPLICANT: Qian, Xiaohong B.

; APPLICANT: Dramac, Radote T.

; TITLE OF INVENTION: No. US20020197679A1el Nucleic Acids and

; FILE REFERENCE: Polypeptides

; CURRENT APPLICATION NUMBER: US/10/098,841

; PRIOR FILING DATE: 2002-03-13

; PRIOR APPLICATION NUMBER: 09/598,042

; PRIOR FILING DATE: 2000-06-20

; PRIOR APPLICATION NUMBER: 09/552,317

; PRIOR FILING DATE: 2000-04-25

; PRIOR APPLICATION NUMBER: 09/488,725

; NUMBER OF SEQ ID NOS: 331

; SOFTWARE: pt_FL_genes Version 1.0

; SEQ ID NO 53

; LENGTH: 1402

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (280)..(1075)

US-10-098-841-53

Alignment Scores:

Pred. No.: 9,84e-110 Length: 1402

Score: 940.00 Matches: 194

Percent Similarity: 74.66% Conservative: 27

Best Local Similarity: 65.54% Mismatches: 41

Query Match: 64.83% Indels: 34

DB: Gaps: 5

US-09-896-522-2 (1-277) x US-10-098-841-53 (1-1402)

QY 2 AlSerAlaGlyGlyIu----- 7

Db 227 GCGGGCGCGGCGGAGCGTGCCTTCGACAGCAGCGGAGGCGGCGCA 286

QY 8 -----AspCyseGluSerProAlaProGluAlaAspArgProHis---GlnArg 22

Db 287 ACATGCGCGGAGACGCGAGACCTTCGACAGACCAACGAGCGGCGGCGAG 346

QY 23 ProPheLeuIleGlyValSerGlyGlyThrAlaSerGlyValSerThrValCysGluVal 42

Db 347 CCTTCCTTATAGCGGTCAAGCGGAGGAGACGTAAGCGGCAAGTCTCCGTGTGCTTAAG 406

QY 43 IleMetGluLeuLeuGlyGlnHisGlnValGlnGlnArgGlnArgValIleLeu 62

Db 407 ATGTGACAGTCTTCGGGCGAGATGAGTGAGTATCGCCAGAGAGGTGTCTATCG 466

QY 63 SerGlnAspArgPheTyrIleValIleuThrAlaGluGlnIleValAlaLeuValGly 82

Db 467 AGCCAGATGAGTCTTACCGTGTCTTACCTCGGAGACAAAGCCCTGAAGGCG 526

QY 527 CAGTTCACTTACACCGGATGCTTGAACATGAGACTCTCTCAAAACACTCAAA 586

QY 103 AsnIleValGlnGlyIleThrValGlnValProThrTyrAspPheValThrHisSerArg 122

Db 587 GAAATCACTGAGAGGAGAAACAGTCCAGATCCCGTATGATGATCTTCCCTCCCG 646

QY 123 LeuProGluThrThrValValTyrProAlaAspValIleuPheGluGlyIleLeuVal 142

Db 647 AAGGAGAGACATTTCTGTCTATCCGACAGAGTGTGCTCTTTAAGGAGACTGCGC 706

QY 143 PheTyrSerGlnIleArgAspMetPheHisLeuArgLeuPheValAspThrAspSer 162

Db 707 TTCTACTCCAGGAGGTAGAGACCTGTTCCAGATGAGCTTTTGTGATACAGATGCG 766

```

QY 163 AspValArcheuSerArgValIleuArgAspVal---ArgArgGlyArgAspLeuGlu 181
DB 767 GACACCGGCTCTCAGCAGAGATTTAAAGGACATCCAGCAGAGAGGAGGATCTTGAG 826
QY 182 GlnIleuThrglnTyrThrThrPheValIysProAlaPheGluGluPheCysLeuPro 201
DB 827 CAGATTTTATCTCAGTACATTAGCTTGTCAAGCTCTCTTGAGGAAATCTGCTTCCA 886
QY 202 ThrIlySerThrValAlaAspValIleIleProArgGlyValAlaPheMetValAlaIleAsn 221
DB 887 ACAAGAAGATGATGTGATGATCATCTCAGAGTCCAGATATATCTGTGGCCATCAAC 946
QY 222 LeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysIlySTpHisArgGly 241
DB 947 CTCATCTGTGAGCAGCATCCAGACATCTTGAAATGA----- 982
QY 242 GlySerAsnGlyArgSerTyrIlyArgThrPheSerGluProGlyAspHisProGlyMet 261
DB 983 -----GGGCCCTCCAAACGGCAGACC-----AATGGCTGTCTCAACGGCTAC 1024
QY 262 LeuThrSerGlyIlyArgSerHisLeuGluSerSerArgProHis 277
DB 1025 ACCCCTTCACGAGAGGAGGAGCATCGAGTCCAGCAGAGCGCGCAT 1072

RESULT 6
US-09-925-300-220
; Sequence 220, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben,
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 220
; LENGTH: 1310
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-925-300-220

Alignment Scores:
Pred. No.: 5,52e-105 Length: 1310
Score: 902.50 Matches: 189
Percent Similarity: 77.82% Conservative: 32
Best Local Similarity: 66.55% Mismatches: 40
Query Match: 62.24% Indels: 24
Gaps: 5

US-09-896-522-2 (1-277) x US-09-925-300-220 (1-1310)
QY 3 SerAlaGlyGlyGluAspCysGluSerProAlaProGluAlaAspArgPro----- 19
DB 205 GCACCGGAGGAGGAGGCGCGCAGCATGCGCGGAGCAGCAGCAGCATCCGCGAGAAC 264
QY 20 HieGlnArg-----ProPheLeuIleGlyValSerGlyIlyThrAlaSer 34
DB 265 CACCAGAGCCCAACGCGCGCGAGCCCTCTTATAGCGCTCAC-GGGGGAACAGCTAGC 323
QY 35 GlySerSerThrValCysGlyIlyIleMetGluLeuLeuGlyGlnAsnGluValGluGln 54
DB 324 GGCAGGCTTCCTGCTGTGCTAGATGTGCAAGCTCTGGGGGAGATGAGGTGACTAT 383
QY 55 ArgGlnArgIlyValAlaIleLeuSerGlnAspArgPheTyrIlyValIleuThrAlaGlu 74
DB 384 CGCAGAGAGGAGGTGTCTGATCGAGCAGATAGCTTCAACGTGTCTTACCTCGAG 443

```

```

QY 75 GlnIlyAlaIlyAlaLeuIlySGlyGlnTyrAsnPheAspHisProAspAlaPheAspAsn 94
DB 444 CAGAGGCGCCAAAGCCCTGAAGGSCAGTTCAACTTTGACCAACCCGAGATGCTTTGACAAAT 503
QY 95 AspLeuMetHisArgThrPheIlyAsnIleValGluGlyIlySerThrValGluValProThr 114
DB 504 GASTGATCTCTCAAAACATCAAAAGAAATCATCGAAGGAGGAAAAACATCCAGATCCCGGTG 563
QY 115 TyrAspPheValThrHisSerArgLeuProGluThrThrValValIlyProAlaAspVal 134
DB 564 TATGACTTGTCTCCCATTCGCCGGAAGAGAGACATTACTGTCTATCCCGCAGACGTG 623
QY 135 ValLeuPheGluGlyIlyLeuValPheTyrSerGlnGluIleArgAspMetPheHisLeu 154
DB 624 GTGCTCTTGAAGAGGATCTCGGCTTCTACTCCAGAGAGATACGAGACCTGTTCCAGAT 683
QY 155 ArgLeuPheValAspThrAspSerAspValArgLeuSerArgArgValLeuArgAspVal 174
DB 684 AACCTTTTGTGATACAGATGGGACACCCGGCTCTCAGCAGATATTAAAGGACATC 743
QY 175 --ArgArgGlyArgAspLeuGluGlnIleLeuThrGlnTyrThrThrPheValIlyPro 193
DB 744 AGCAGAGAGGCGAGGATCTTGACAGATTTTATCTCAGTACATTCAGTTCCGCAAGCCT 803
QY 194 AlaPheGluGluPheCysLeuProThrIlyIlySerThrAlaAspValIleIleProArgGly 213
DB 804 GCTTTGAGAAATTCGTCTGCTCCCAACAAAGATATGTGATGTATCATCTCTTAAGGT 863
QY 214 ValAspAspMetValAlaIleLeuLeuIleValGlnHisIleGlnAspIleLeuAsnGly 233
DB 864 GCAGATATCTGTGGCCATCAACCTCATCTGTCAGACATCCAGACATCTCGAATGA 923
QY 234 AspIleCysIlySTpHisArgGlyIlySerAsnGlyArgSerTyrIlyArgThrPheSer 253
DB 924 -----GGGCCCTCCAAACGGCAGACC----- 944
QY 254 GluProGlyAspHisProGlyMetLeuThrSerGlyIlyArgSerHisLeuGluSerSer 273
DB 945 --AATGGCTGTCTCAACGGCTTACACCTCTCAGCAGAGGAGGAGCATCGAGTCCAGC 1001
QY 274 SerArgProHis 277
DB 1002 AGCAGGCGCGCAT 1013

RESULT 7
US-09-918-995-30379
; Sequence 30379, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 30379
; LENGTH: 472
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)---(472)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-30379

Alignment Scores:
Pred. No.: 3.18e-83 Length: 472
Score: 727.00 Matches: 143

```

Percent Similarity: 99.31% Conservative: 0
 Best Local Similarity: 99.31% Mismatches: 1
 Query Match: 50.14% Indels: 0
 DB: 11 Gaps: 0

US-09-896-522-2 (1-277) x US-09-918-995-30379 (1-472)

QY 33 A1aSerGlyLysSerThrValCysGluLysIleMetGluLeuGluValAsnGluVal 52
 Db 41 GCCACGGGAAAGTCAACCTGTGTGAGAAATCATGAGTTGCTGGACAGAGAGGTG 100
 QY 53 GluGlnArgGlnArgLysValValIleLeuSerGlnAspArgPheTyrLysValLeuThr 72
 Db 101 GAACAGCGCGACGGAAGGTGATCTCTGAGCCGACAGATCTTACAAAGTCTTGACG 160
 QY 73 A1aGluGlnLysAlaLysAlaLeuLysGlyLysIleTyrAsnPheAspHisProAspAlaPhe 92
 Db 161 GCAGAGCAGAAAGGCCMAAGCCTTGAAAGACAGTATTTGACATCCAGATGCCCTTT 220
 QY 93 AspAsnAspLeuMetHisArgThrLeuLysAsnIleValGluGlyThrValGluVal 112
 Db 221 GATATGATTTGATGACACAGACTCTGAAGACATCGTGAAGGCAAAACGTTGAGGTG 280
 QY 113 ProThrTyrAspPheValThrHisSerArgLeuProGluThrThrValValTyrProAla 132
 Db 281 CCGACCTATGATTTGTGTACACACTCAAGGTTTACAGAGACCAAGGTGCTTACCTGCG 340
 QY 133 AspValValIleuPheGluGlyIleLeuValPheTyrSerGlnGluIleArgAspMetPhe 152
 Db 341 GACGTGTTCTGTTTGAAGGCGATCTGTGTTCTTACAGCCAGAGATCCGGGACATGTTCC 400
 QY 153 HisLeuArgLeuPheValAspThrAspSerAspValArgLeuSerArgArgValLeuArg 172
 Db 401 CACCTGCGCCCTTCTGTGTGACACACGACTCCGACGCTGATCGAAGGTTCTCCGG 460
 QY 173 AspValArgArg 176
 Db 461 GACGTGCGCCGA 472

RESULT 8

US-09-764-877-2805
 ; Sequence 2805, Application US/09764877
 ; Patent NO. US20020147140A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosen et al.
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
 ; FILE REFERENCE: PC005
 ; CURRENT APPLICATION NUMBER: US/09/764,877
 ; PRIORITY FILING DATE: 2001-01-17
 ; Prior application data removed - refer to PALM or file wrapper
 ; NUMBER OF SEQ ID NOS: 4031
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 2805
 ; LENGTH: 19125
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-09-764-877-2805

Alignment Scores:

Pred. No.: 1,53e-70 Length: 19125
 Score: 647.00 Matches: 197
 Percent Similarity: 26.51% Conservative: 1
 Best Local Similarity: 26.37% Mismatches: 3
 Query Match: 44.62% Indels: 548
 DB: 10 Gaps: 4

US-09-896-522-2 (1-277) x US-09-764-877-2805 (1-19125)

QY 1 MetAlaSerAlaGlyGlyGluAspCysGlnSerProAlaProGluAlaAspArgProHis 20
 Db 41 ATGGCTTCGGCGGAGCGGAGACTCGAGAGCCCGCGCGGCGGAGCCGCTCCGAC 100
 QY 21 GlnArgProPheLeuIleGlyValSerGlyGlyThrAlaSerGlyLys----- 36

Db 101 CAGGCGCCTTCTTGATAGGGGTGAGGCGGCGGCACTGCGAGCGGGAAGTAAAGGGCCGA 160
 QY 36 ----- 36
 Db 161 GCGGCGCGGGTCTCGCTCTCTCGCGGTGCTTAAGGCGGCGCCGGTCCAGAGAGAGACC 220
 QY 36 ----- 36
 Db 221 CCAAGCCCGCCAGAGGGGCCCTTCCAGGACAGGGGCGAGGGCGGGAAGGGGCTG 280
 QY 36 ----- 36
 Db 281 CAGGCGCGATGGGCGGCGGGGCTTGGGCGGCGAGGGGTTGGGGCGCTGGGCGCTCG 340
 QY 36 ----- 36
 Db 341 GGCTCCCGCAGACCTCTCCGCGGCGGCTCTCGAGCCCAAGTCCGTCTAAACC 400
 QY 36 ----- 36
 Db 401 CGTTGGGCGCCCGGCGCCCTTCCGTGAGACCATCAACACTGTGTGCTCTCCGCGCT 460
 QY 36 ----- 36
 Db 461 TCGAGCATACCCCCAGGGGGGGGGTGGGTCAGTGTTCCAGAACCGGCGCGTGGCT 520
 QY 37 -----SerThrValCys 40
 Db 521 GTCAAGAGCAGACCGGTTCCAGCGGGAAATCCCGCTGTGTCTTCAGTCGACCTGTG 580
 QY 40 sGluLysIleMetGluLeuLeuGlyGlnAsnGluValGluGlnArgGlnArgLysValVa 60
 Db 581 TGACAAATATATGAAGTTGCTGTGGAACAGACGAGTGAACAGCCGCGAGGAGGTGT 640
 QY 60 lIleLeuSerGlnAspArgPheTyrLysValLeuThrAlaGluGlnLysAlaLysAla 80
 Db 641 CATCTGAGCCAGACGAGGTTCTTACAAAGTCTGAGCGGACAGACAGCCAGGCCACTT 700
 QY 80 uLysGlyGlnTyrAsnPheAspHisPro----- 88
 Db 701 GAAAGGACAGTACAAATTTTGAACATCCAGGTAATCGCGAGCCCAACGATGGGAGAG 760
 QY 89 ----- 89
 Db 761 CAGACCCCATTTCTGGCGCGCCGCTAACAGCTGCGCTTGGGCCCAAGTTGATGATGC 820
 QY 89 ----- 89
 Db 821 AGACCGGCTTGGGGTGTTCCTCCCAACTGCGCCAGCGAGCTGCAGTGAGTGGCTGCA 880
 QY 89 ----- 89
 Db 881 CAGTGAAGTCAGGGAACCTTGTGAGTTCTCTGGGCCCTGAGTGTCCCCCAGAGCGT 940
 QY 89 ----- 89
 Db 941 GGTGAGGACAGTCAGAGAAAGCTGTTCACACTGTGGGGAAGCCCTGCTGCTGCCACA 1000
 QY 89 ----- 89
 Db 1001 CTGTGATCTGAGAAATCTCCCATCTTCCCTTCCAGAGTTGGGGCTTTAGGGCGCTC 1060
 QY 89 ----- 89
 Db 1061 CACACACACTCTGCCACTGTGCTCTCCGCTTGGGGGCAAAAGTCTCAGAGAGGCC 1120
 QY 89 ----- 89
 Db 1121 AGTGTGGAGGCGGTAGTAACTAGTGTGCTGTGAGCTTGAAGCTCCGGGCGAGGTAAAG 1180
 QY 89 ----- 89

Db 1181 GGCAGATGCGGAGGAGACACCTGCACTGCTAGAGATTCCACCTGGCTGGCTGGCA 1240
 Qy 89 ----- 89
 Db 1241 TGGTCAGCCCTTTTCCCAAGGACTTTGGGAGAAAGCGCTGCAAGCCGAGCTTAACGAG 1300
 Qy 89 ----- 89
 Db 1301 TTAGTCCCAAGAAAG 1360
 Qy 89 ----- 89
 Db 1361 TCAGAAAACCCCTGTGTTCGGGATCTCAGGGGACAGAGCTTGGGAGTGGACAGAGCC 1420
 Qy 89 ----- 89
 Db 1421 GAGAACAGCAAGAGACAGGCTATCCTGGTCTGACAGGCTGACAGAGGACTTGGCTA 1480
 Qy 89 ----- 89
 Db 1481 CCAAACTACACTGCGCTAGTCTAGTCTAGTCTAGGCTTCTGGTGTCCCAAGCA 1540
 Qy 89 ----- 89
 Db 1541 CCCAGCTGTGCTGACCCCAAGAGAGAGAGGTTCCCTGTGCTGCTGCAAGAGGCTATC 1600
 Qy 90 ----- 90
 Db 1601 TGCCTGTCTTGCCTTTGTCTCCAACTGCAAGATGCTTTGATTAATTAATTAATTAATTA 1660
 Qy 99 GTTtLeuLyAsn11eValGluGlySerThr-ValGluValProThrTyrAspPheValT 119
 Db 1661 GACTCTGAACACCATCTGTGAGAGGCAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1720
 Qy 119 hrHisSer----- 121
 Db 1721 CACACTC-AAGGTAAAGAGAGTGTGTTTAAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1779
 Qy 121 ----- 121
 Db 1780 GTCCCTCGGGCTCACACAGACTGCTGCGGAGGCTCCCGGCTTGTGTTCTGTGG 1839
 Qy 121 ----- 121
 Db 1840 GTCCCCACCCCTTCCACAGTCACTGCTGCTGCAATGAACTGACTGTGTGTGGCC 1899
 Qy 122 ----- 126
 Db 1900 TGGAGACCTGTGTCTGACAGAGAGCCCTTCCCGGTGGGCGCCCTCAGGTTTACAGAGA 1959
 Qy 126 hrThrValValTyrProAlaAspValValLeuPheGluGlyIleLeuValPheTyrSerG 146
 Db 1960 CCAAGGAGGCTACCTCCGAGAGGAGTGTGTTGTTGAGGAGCATCTTGTCTTACAGCC 2019
 Qy 146 lngluIleArgAspMetPheHisLeuArgPhePheValAspThrAspSerAspValArgL 166
 Db 2020 AGAGAGATCCGGAGCATCTTCACTGCGCTTTCGTGACACCGACTCCGAGCTGAGGC 2079
 Qy 166 euSerArgArgVal----- 170
 Db 2080 TGTCTCGAAGAGG-TAAGGCGCGCGCGGCTCCGTGCTGCTGCGCGAGGCTGGCGGC 2138
 Qy 171 ----- 176
 Db 2139 AGGTCTGAGGTGTGAAGACAGCGCTGCGCTTGGCGCTGAGTCTCCCGGAGAGTGGGCC 2198
 Qy 176 rgglYArgAspLeuGluGlnIleLeuThrGlnTyrThrThrPheValValProAlaPheG 196
 Db 2199 GAGGAGGAGGAGCTGAGGAGATTTCTGACGAGTACACCACTTCTGTGAAGCCGGCTTGC 2258
 Qy 196 lngluPheCysLeuPro 201
 Db 2259 AGGAGTTCTGCTGCGG 2275

RESULT 9
 US-09-938-842A-1943
 ; Sequence 1943, Application US/09938842A
 ; Patent No. US20020160378A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Harper, Jeff
 ; APPLICANT: Kreps, Joel
 ; APPLICANT: Wang, Xun
 ; APPLICANT: Zhu, Tong
 ; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
 ; FILE OF INVENTION: SAME, AND METHODS OF USE
 ; FILE REFERENCE: SRIPI300-3
 ; CURRENT APPLICATION NUMBER: US/09/938,842A
 ; CURRENT FILING DATE: 2001-08-24
 ; PRIOR APPLICATION NUMBER: US 60/227,866
 ; PRIOR FILING DATE: 2000-08-24
 ; PRIOR APPLICATION NUMBER: US 60/264,647
 ; PRIOR FILING DATE: 2001-01-16
 ; PRIOR APPLICATION NUMBER: US 60/300,111
 ; PRIOR FILING DATE: 2001-06-22
 ; NUMBER OF SEQ ID NOS: 5379
 ; SEQ ID NO 1943
 ; LENGTH: 1401
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana
 US-09-938-842A-1943
 Alignment Scores:
 Pred. No.: 4,166-59 Length: 1401
 Score: 544.00 Matches: 116
 Percent Similarity: 63.16 Conservative: 52
 Best Local Similarity: 43.618 Mismatches: 78
 Query Match: 37,528 Indels: 20
 Gaps: 8
 US-09-896-522-2 (1-277) x US-09-938-842A-1943 (1-1401)
 Qy 11 SerProAlaProGlu-----AlaSerArgProHisGlnArgProPheLeuIleGly 27
 Db 85 TCMAATAGACACAGAGCAATGCTGAG 144
 Qy 28 ValSerGlyGlyThrValSerGlyLysSerThrValCysGluLysIleMetGluLeuLeu 47
 Db 145 GTTGTGAGGTGACACATCCGGAGAAACACTGTGTGTGATGATGATGATGATGATGATGATGATGAT 204
 Qy 48 GlyGlnAsnGluValGluGlnArgGlnArgLysValValIleLeuSerGlnAspArgPhe 67
 Db 205 -----CATGATCAGAGAGCTGTTGTGTAATCAGGATTCCTTC 243
 Qy 68 TyrLysValLeuThrAlaGluGlnLysValAlaLeuLysGlyGlnTyrAsnPheAsp 87
 Db 244 TTCACATTAATGTAATGTAATGAGCTGTGTAAGAGTT-----CATGATTAACAATTTTGAC 297
 Qy 88 HisProAspAlaPheAspAsnAspLeuMetHisArgThrLeuLysAsnIleValGluGly 107
 Db 298 CATCTGACGCTTTCATCTGACATCTGACATTAATTTCTTCATGAGAGAGTTTAAGAAAGAG 357
 Qy 108 LysThrValGluValProThrTyrAspPheValThrHisSerArg-----LeuProGlu 125
 Db 358 CAAGCAGTGAATATCTTAATTAATGACTTCAAAAGTTACAAAGAACAAATGTTTTCACACT 417
 Qy 126 ThrThrValValTyrProAlaAspValValLeuPheGluGlyIleLeuValPheTyrSer 145
 Db 418 AGAAGG---GTGAATCTTCTGACGTTAATTAATTCGGAAGGTAACTCAATTTCCATGAC 474
 Qy 146 GlnGluIleArgAspMetPheHisLeuArgLeuPheValAspThrAspSerAspValArg 165
 Db 475 CTCGGGTCGAGATTTGTAAGACATGAAGATATTTGTGAAGCAGATCCCATGTGCT 534
 Qy 166 LeuSerArgArgValLeuArgAsp---ValArgArgGlyArgAspLeuGlnIleLeu 184
 Db 535 CTAGCGAAGAAAGATTAAGCGCATACTGTTGAAGAGGAGAGAGATTAATCCACTGTCTT 594

Qy 185 ThrGlnTyrThrPheValIysProAlaPheGlnGluPheCysIleuProThrLysIys 204
 Db 595 GACGAGTCTCAAGATTGTGTGAAGCCAGCATTTGAGATTTCATATCTCCACAAAGAA 654
 Qy 205 TyrAlaAspValIleIleProArgIyValAspAsnMetValAlaIleAsnIleVal 224
 Db 655 TACGAGATATATATATATCCCGAGGTGTGTATTAACATGTTGCTATTGATTGTG 714
 Qy 225 GlnHisIleGlnAspIleLeu---AsnGlyAspIleCysIlySTPPhIAsrGlyIys 243
 Db 715 CAACACATCCATACAGAGCTGGTCAACATGATCTGTAAATTAT-----CCA 765
 Qy 244 AsnGlyAsrSerIyTyrArgThrPheSerGluProGlyAspHisProGlyMetIleThr 263
 Db 766 AATCTTATGTATTATTCATCACTTTTCAGATACGTGTATGACACTCTATATCCGAC 825
 Qy 264 SerGlyAsrSerHis 269
 Db 826 TCTAAACACAAAGCAT 843

RESULT 10

US-10-251-186-14
 ; Sequence 14, Application US/10251186
 ; Publication No. US20030180745A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Tang, Y. Tom
 ; APPLICANT: Zhang, Jie
 ; APPLICANT: Drmanac, Radoje T.
 ; TITLE OF INVENTION: No. US20030180745A1el Nucleic Acids and
 ; FILE REFERENCE: 789CIP2CDV1
 ; CURRENT APPLICATION NUMBER: US/10/251,186
 ; PRIOR FILING DATE: 2002-09-19
 ; PRIOR APPLICATION NUMBER: 09/665,363
 ; PRIOR FILING DATE: 2000-09-19
 ; PRIOR APPLICATION NUMBER: 09/574,454
 ; PRIOR FILING DATE: 2000-05-19
 ; PRIOR APPLICATION NUMBER: 09/519,705
 ; PRIOR FILING DATE: 2000-03-07
 ; NUMBER OF SEQ ID NOS: 35
 ; SOFTWARE: pc_FL_genes Version 2.0
 ; SEQ ID NO 14
 ; LENGTH: 2058
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (173)..(1909)
 US-10-251-186-14

Alignment Scores:

Pred. No.: 2,1e-58 Length: 2058
 Score: 540.50 Matches: 110
 Percent Similarity: 64.20% Conservative: 46
 Best Local Similarity: 45.27% Mismatches: 60
 Query Match: 37.28% Indels: 27
 DB: 12 Gaps: 5

US-09-896-522-2 (1-277) x US-10-251-186-14 (1-2058)

Qy 9 CysGluSerProAlaPro-----GluAlaAspArg 18
 Db 455 TGCAGTCAAGAGCTCCCTCTGCTGCTACAGCAAGCTACATCTACACCGCGCGG 514
 Qy 19 Pro-----HisGlnArgProPheLeuIleGlyValSer 29
 Db 515 CCGGCTGTGTAACATGAACAGCAGCAGCAATCAAGAGGCTTCGCGCATGCGTTGGA 574
 Qy 30 GlyGlyThrAlaSerGlyIysSerThrValCysGluIysIleMetGluLeuGln 49
 Db 575 GCGCGAGTGTCTCTGGGAAACCACTGTGCGCAGAAATGATCATCGAGCGCTG----- 628

Qy 50 AsnGluValGlnGluArgGlnArgIysValIleIleuSerGlnAspArgPheIyIys 69
 Db 629 ---GATGGCCCTGG-----GTGGTCTTGCTGCTCATGAGCTCTTCTACAG 673
 Qy 70 ValIleuThrAlaGlnGlnIyValAlaIleuIyGlyGlnIyTrAsnPheAspHisPro 89
 Db 674 GTGCTGACTGAGCAGCAGGAGGAGGAGGCGCAGACAACTCAACTCACTCGACACCA 723
 Qy 90 AspAlaPheAspAsnIleuMetHisArgThrIleuIyAsnIleValGlnGlyIyThr 109
 Db 734 GATGCTTTGACTTCGACCTTCATCATTTCCACCTCAAGAGGCTGAAGCGGGAAGGT 793
 Qy 110 ValGluValProThrTyrAspPheValThrHisSerArgIleuProGluThrThrVal 129
 Db 794 GTCAAGGTGCCCTTTATGACTTCACACCGCAGCGGAGAGACTGGAACACTG 853
 Qy 130 TyrProAlaAspValValIleuPheGluGlyIleLeuValPheTyrSerGlnIleArg 149
 Db 854 TATGCTGCAAAACGTCATCATCTTGAAGGCATCATGCGCTTGTCTGACAAAGACCTGTTG 913
 Qy 150 AspMetPheHisIleuArgIleuPheValAspThrAspSerAspValArgIleuSerArg 169
 Db 914 GAGCTCCTGGACATGAAGATCTTTGTGACACAGACTCCGACATCCGCTGTACGCGG 973
 Qy 170 ValIleuArgAspVal---ArgArgGlyArgAspIleuGlnIleuThrGlnTyrThr 188
 Db 974 CTGCGCCGGGACATCACTGAGCCCGCGGAGCATGAGGCTCATCAAGCATCAAC 1033
 Qy 189 ThrPheValIysProAlaPheGlnGluPheCysLeuProThrLysIyTyrAlaAspVal 208
 Db 1034 AAGTTGTCAAGCCTCTTCGACAGCATCACTCAACCCCATGCGCGCTGGCAGACATC 1093
 Qy 209 IleIleProArgIyValAspAsnMetValAlaIleAsnIleValGlnHisIleGln 228
 Db 1094 GTGGTCCCAAGAGGAGCGGAAACGCGTGCATGACCTGATGTGGACGCTGCAC 1153
 Qy 229 AspIleLeu 231
 Db 1154 AGCCAGCTG 1162

RESULT 11

US-09-918-995-24042
 ; Sequence 24042, Application US/09918995
 ; Publication No. US20030073623A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Hyseq, Inc.
 ; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
 ; FILE REFERENCE: 20411-756
 ; CURRENT APPLICATION NUMBER: US/09/918,995
 ; PRIOR FILING DATE: 2001-07-30
 ; PRIOR APPLICATION NUMBER: US/09/235,076
 ; PRIOR FILING DATE: 1999-01-20
 ; NUMBER OF SEQ ID NOS: 38054
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 24042
 ; LENGTH: 472
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (1)...(472)
 ; OTHER INFORMATION: n = A, T, C or G
 US-09-918-995-24042

Alignment Scores:

Pred. No.: 5.36e-59 Length: 472
 Score: 537.50 Matches: 103
 Percent Similarity: 84.14% Conservative: 19
 Best Local Similarity: 71.03% Mismatches: 22
 Query Match: 37.07% Indels: 1
 DB: 11 Gaps: 1

US-09-896-522-2 (1-277) x US-09-918-995-24042 (1-472)

Qy 8 AspCysGluSerProAlaProGluAlaAspArgProHis---GlnArgProPheLeuIle 26
 Db 37 GACAGCAGACGACACCTGCGAAGACCACCGAGCCGCGGCGAGCCCTTCTTATA 96
 Qy 27 G1YValSerG1Yg1YThrAlaSerG1YlysSerThrValCysG1uYsIleMetG1uLeu 46
 Db 97 GGCCTCAGCGGGGGAACAGCTACGCGCAAGCTTCCTGTGTCTTAAGATCGTCAGCTC 156
 Qy 47 LeuG1Yg1uAsnG1uValG1uGlnArgG1uArgIysValVal11leuSerG1uAspArg 66
 Db 157 CTGGGGCAGATAGTGAGTGAATCGCCAGACGAGTGTCTCATCTGAGCCAGATAGC 216
 Qy 67 PheTyrIysValLeuThrAlaGluGlnIlysa1AlaYsa1AlaYsa1G1uYsIleValAspPhe 86
 Db 217 TTCTACCGGTCTCTTACTCTGAGCAGAAAGCCCAAGCCCTGAAGGCGCAGTTCACTTT 276
 Qy 87 AsphiSProAspAlaPheAspAsnAspLeuMetHisArgThrLeuYsaen1IleValG1u 106
 Db 277 GACCAACCGGATGCTTGTGACATGAATCATCTCTCAAAACACTCAAAAGAAATCACTGAA 336
 Qy 107 G1YlysThrValG1uValProThrTyrAspPheValThriSserArgLeuProGluThr 126
 Db 337 GGGAAAACAGTCCAGATCCCGGTATGACTTGTCTCCCATTCGCGAAGAGAGACA 396
 Qy 127 ThrValValTyrProAlaAspValValLeuPheGluG1Y1leuValPheTyrSerGln 146
 Db 397 GTTACTGTCTATCCCGCAGAGAGTGCTCTTTGAAGGAGATCCGCGCTTCTACTCCAG 456
 Qy 147 Glu1leArgAspMet 151
 Db 457 GAGGTACGAGACCTG 471

RESULT 12

US-10-066-543-1843
 ; Sequence 1843, Application US/1006543
 ; Publication No. US20030087818A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Jiang, Yuguo
 ; APPLICANT: Pye, Ruth A.
 ; APPLICANT: Xu, Jiangchun
 ; APPLICANT: Indrias, Carol Yoseph
 ; APPLICANT: Lodes, Michael J.
 ; APPLICANT: Secrist, Heather
 ; APPLICANT: Carter, Darick
 ; APPLICANT: Fanger, Gary R.
 ; APPLICANT: Smith, Carole L.
 ; APPLICANT: Durham, Margarita
 ; APPLICANT: Stolk, John A.
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
 ; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
 ; FILE REFERENCE: 210121.563
 ; CURRENT APPLICATION NUMBER: US/10/066,543
 ; NUMBER OF SEQ ID NOS: 3417
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 1843
 ; LENGTH: 447
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-10-066-543-1843

Alignment Scores:

Pred. No.: 5,47e-57 Length: 447
 Score: 521.50 Matches: 101
 Percent Similarity: 83.69% Conservative: 17
 Best Local Similarity: 71.63% Mismatches: 22
 Query Match: 35.97% Indels: 1
 DB: 14 Gaps: 1

US-09-896-522-2 (1-277) x US-10-066-543-1843 (1-447)

Qy 8 AspCysGluSerProAlaProGluAlaAspArgProHis---GlnArgProPheLeuIle 26
 Db 23 GACAGCAGACGACACCTGCGAAGACCACCGAGCCGCGGCGAGCCCTTCTTATA 82
 Qy 27 G1YValSerG1Yg1YThrAlaSerG1YlysSerThrValCysG1uYsIleMetG1uLeu 46
 Db 83 GGCCTCAGCGGGGGAACAGCTACGCGCAAGCTTCCTGTGTCTTAAGATCGTCAGCTC 142
 Qy 47 LeuG1Yg1uAsnG1uValG1uGlnArgG1uArgIysValVal11leuSerG1uAspArg 66
 Db 143 CTGGGGCAGATAGTGAGTGAATCGCCAGACGAGTGTCTCATCTGAGCCAGATAGC 202
 Qy 67 PheTyrIysValLeuThrAlaGluGlnIlysa1AlaYsa1AlaYsa1G1uYsIleValAspPhe 86
 Db 203 TTCTACCGGTCTCTTACTCTGAGCAGAAAGCCCTGAAGGCGCAGTTCACTTT 262
 Qy 87 AsphiSProAspAlaPheAspAsnAspLeuMetHisArgThrLeuYsaen1IleValG1u 106
 Db 263 GACCAACCGGATGCTTGTGACATGAATCATCTCTCAAAACACTCAAAAGAAATCACTGAA 322
 Qy 107 G1YlysThrValG1uValProThrTyrAspPheValThriSserArgLeuProGluThr 126
 Db 323 GGGAAAACAGTCCAGATCCCGGTATGACTTGTCTCCCATTCGCGAAGAGAGACA 382
 Qy 127 ThrValValTyrProAlaAspValValLeuPheGluG1Y1leuValPheTyrSerGln 146
 Db 383 GTTACTGTCTATCCCGCAGAGAGTGCTCTTTGAAGGAGATCCGCGCTTCTACTCCAG 442
 Qy 147 Glu 147
 Db 443 GAG 445

RESULT 13

US-09-918-995-23923
 ; Sequence 23923, Application US/09918995
 ; Publication No. US20030073623A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Hyseq, Inc.
 ; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
 ; TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
 ; FILE REFERENCE: 20411-756
 ; CURRENT APPLICATION NUMBER: US/09/918,995
 ; CURRENT FILING DATE: 2001-07-30
 ; PRIOR APPLICATION NUMBER: US/09/235,076
 ; PRIOR FILING DATE: 1999-01-20
 ; NUMBER OF SEQ ID NOS: 38054
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 23923
 ; LENGTH: 455
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc.feature
 ; LOCATION: (1)...(455)
 ; OTHER INFORMATION: n = A,T,C or G
 ; US-09-918-995-23923

Alignment Scores:

Pred. No.: 1,07e-55 Length: 455
 Score: 511.50 Matches: 99
 Percent Similarity: 83.45% Conservative: 17
 Best Local Similarity: 71.22% Mismatches: 22
 Query Match: 35.28% Indels: 1
 DB: 11 Gaps: 1

US-09-896-522-2 (1-277) x US-09-918-995-23923 (1-455)

Qy 8 AspCysGluSerProAlaProGluAlaAspArgProHis---GlnArgProPheLeuIle 26
 Db 37 GACAGCAGACGACACCTGCGAAGACCACCGAGCCGCGGCGAGCCCTTCTTATA 96
 Qy 27 G1YValSerG1Yg1YThrAlaSerG1YlysSerThrValCysG1uYsIleMetG1uLeu 46
 Db 27 G1YValSerG1Yg1YThrAlaSerG1YlysSerThrValCysG1uYsIleMetG1uLeu 46

Db 97 GGCCTCAGCGGGGAGACAGTACGCGCAAGTCTTCCTCGTGTGCTAAGATCGTCAGCTC 156
Qy 47 LeuGlyGlnAsnGluValGluGlnArgGlnArgValValIleLeuSerGlnAspArg 66
Db 157 CTGGGCGAATGAGTGAAGTACTATCGCAGAGCAGGCGTGCATCTTACAGCAGATAGC 216
Qy PheTyrIysValIleThrAlaGluGlnIysAlaIysAlaLeuIysGluGlnTyrAspPhe 86
Db 217 TTTCACCTGTCTTACTCTCGAGAGAGAGCCAAAGCCCTGMAAGCCATTCTACTT 276
Qy 87 AspHisProAspAlaPheAspAsnAspLeuMetHisArgThrLeuIysAsnIleValGlu 106
Db 277 GACCAACCGGATGCTTGAACAATGACTCATCTTCAAAACACTCAAGAATACATGAA 336
Qy 107 GlyIysThrValGluValProThrTyrAspPheValThrHisSerArgLeuProGluThr 126
Db 337 GGGAAAGATCGATCCCGGTATGACTTGTCTCCATCCCGAGAGAGAGACA 396
Qy 127 ThrValValTyrProAlaAspValValLeuPheGluGluIleLeuValPheTyrSer 145
Db 397 GTTACTGTCTATCCCGCAGACGTGTGCTCTTTGAAGGATCTTGCCCTTACTCC 453

RESULT 14
US-09-070-927A-215/c
Sequence 215, Application US/09070927A
Patent No. US20020120116A1
GENERAL INFORMATION:
APPLICANT: Charles A. Kunach
Patrick J. Dillon
Steven Barash
TITLE OF INVENTION: Enterococcus faecialis Polynucleotides and Polypeptides
NUMBER OF SEQUENCES: 982
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/070,927A
FILING DATE: 04-May-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/046,655
FILING DATE: 1997-05-16
APPLICATION NUMBER: 60/044,031
FILING DATE: 1997-05-06
APPLICATION NUMBER: 60/066,009
FILING DATE: 1997-11-14
ATTORNEY/AGENT INFORMATION:
NAME: Kenley K. Hoover
REGISTRATION NUMBER: 40,302
REFERENCE/DOCKET NUMBER: PB369
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 215:
SEQUENCE CHARACTERISTICS:
LENGTH: 3662 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 215:
US-09-070-927A-215
Alignment Scores: 1.64e-43 Length: 3662
Pred. No.:

Score: 427.00 Matches: 82
Percent Similarity: 62.32% Conservative: 52
Best Local Similarity: 38.32% Mismatches: 64
Query Match: 29.45% Indels: 16
DB: 10 Gaps: 4

US-09-896-522-2 (1-277) x US-09-070-927A-215 (1-3662)

Qy 22 ArgProPheLeuIleGlyValSerGlyIleThrAlaSerGlyIysSerThrValCysGlu 41
Db 3328 CAGCAATTTCATTCGTGTACTGGCGGATCGAAGGCGGAAGAACCGATGACCGCT 3269
Qy 42 LysIleMetGluLeuLeuGlyGlnAsnGluValGluGlnArgValIle 61
Db 3268 GCGATTTCATTAATTTCCGATCAT-----TCAATTATGATG 3230
Qy 62 LeuSerGlnAspArgPheTyrIys-----ValLeuThrAlaGluIleValIys 78
Db 3229 TTGAACACGATTCCTATATTAATAAGATCAGAGCCATTGAGTTTGAAGAACCTTAAAT 3170
Qy 79 AlaLeuIysGlyIleTyrAsnPheAspHisProAspAlaPheAspAsnAspLeuMetHis 98
Db 3169 ACC-----AATTATGATCATCTTTGCGTTGATACAGATTAATTGATT 3125
Qy 99 ArgThrLeuIysAsnIleValGluGlyIysThrValGluValProThrTyrAspPheVal 118
Db 3124 CAACATGTTGAACAACGTTAAATTAATCAAGCCATGAAAGCCAGTTATGACTACGTT 3065
Qy 119 ThrHisSerArgLeuProGluThrThrValValTyrProAlaAspValValLeuPheGlu 138
Db 3064 GCGCATTCAGAAAGTACAGAAACGATGATCAAGAACCGGAAGAAAGTAATTAATCTTGAA 3005
Qy 139 GlyIleLeuValPheTyrSerGlnGluIleArgAspMetPheHisLeuArgLeuPheVal 158
Db 3004 GGGATTTCATTTAGAGATGCGCCGCTTAAGGATTTAATGATTTAAAGGTATGTG 2945
Qy 159 AspThrAspSerAspValArgLeuSerArgArgValLeuArgAspVal--ArgArgGly 177
Db 2944 GATACGATGATGATCGATTCGATTAATTCGTCGATCAAGCGTATATGAAAGCGTGGC 2885
Qy 178 ArgAspLeuGluGlnIleLeuThrGlnTyrThrThrPheValIysProAlaPheGluGlu 197
Db 2884 CGTACGTTGATTCGTGATGTAACATATTTGACAGCTCCTCAAGCCCATGATTCACCA 2825
Qy 198 PheCysLeuProThrIysIysTyrAlaAspValIleIleProArgGlyValAspAsnMet 217
Db 2824 TTATTGACCTTCGAAACGTTATGCAATATCATTTGTCCAGAAAGTGGCGAAACACAC 2765
Qy 218 ValAlaIleAsnLeuIleValGlnHisIleGlnAspIleLeu 231
Db 2764 GTCGCTATGATTTAATTAATACAAAGTGATAGATTTTA 2723
RESULT 15
US-10-329-960-1
Sequence 1, Application US/10329960
Publication No. US20030099277A1
GENERAL INFORMATION:
APPLICANT: Fleischmann et al.
TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus influenzae Rd Genome, Fragm
FILE REFERENCE: PB186P1
CURRENT APPLICATION NUMBER: US/10/329,960
CURRENT FILING DATE: 2003-01-02
PRIOR APPLICATION NUMBER: US 09/643,990
PRIOR FILING DATE: 2000-08-23
PRIOR APPLICATION NUMBER: US 08/487,429
PRIOR FILING DATE: 1995-06-07
PRIOR APPLICATION NUMBER: US 08/426,787
PRIOR FILING DATE: 1995-04-21
NUMBER OF SEQ ID NOS: 1
SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
LENGTH: 1830121

```
TYPE: DNA
ORGANISM: Haemophilus influenzae
FEATURE:
NAME/KEY: misc_feature
LOCATION: (4747)..(4747)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (9921)..(9921)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (10150)..(10150)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (29298)..(29298)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (36543)..(36543)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (36551)..(36551)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (36636)..(36636)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (40808)..(40810)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (44416)..(44416)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (44905)..(44905)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (44975)..(44975)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (45593)..(45593)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (45732)..(45732)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (47036)..(47036)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51334)..(51334)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51602)..(51602)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51786)..(51786)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51805)..(51805)
OTHER INFORMATION: n equals a, t, g or c

OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (55369)..(55369)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (65309)..(65309)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (65313)..(65313)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (80024)..(80024)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (100091)..(100091)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (102696)..(102696)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (105121)..(105121)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (107248)..(107248)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (117136)..(117136)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (119750)..(119750)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (119924)..(119924)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (120038)..(120038)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (121344)..(121344)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122167)..(122167)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122336)..(122336)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131340)..(131340)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131360)..(131360)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (139910)..(139910)
OTHER INFORMATION: n equals a, t, g or c
```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: November 25, 2003, 08:09:36 ; Search time 317 Seconds

(without alignments)
2358.815 Million cell updates/sec

Title: US-09-896-522-2

Perfect score: 1450

Sequence: 1 MASAGEDCSPPAEADRP.....HPGMUTSGKSHLESSSRH 277

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+.p2n.model -DEV=x1h
-Q=/c9n2_1/USPTO.spool/US09896522/runcat_21112003_184353_4997/app_query.fasta_1.455
-DB=N Geneseq_19Jun03 -OPMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPEXT=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=bloms62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=p2n -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09896522.OCGN_1.1.312.0/runcat_21112003_184353_4997 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARM TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N_Geneseq_19Jun03.*
1: /SIDSI/gcgdata/geneseg/geneseqn-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseg/geneseqn-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseg/geneseqn-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseg/geneseqn-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseg/geneseqn-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseg/geneseqn-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseg/geneseqn-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseg/geneseqn-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseg/geneseqn-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseg/geneseqn-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseg/geneseqn-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseg/geneseqn-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseg/geneseqn-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseg/geneseqn-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseg/geneseqn-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseg/geneseqn-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseg/geneseqn-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseg/geneseqn-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseg/geneseqn-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseg/geneseqn-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseg/geneseqn-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseg/geneseqn-emb1/NA2001.DAT.*
23: /SIDSI/gcgdata/geneseg/geneseqn-emb1/NA2001B.DAT.*
24: /SIDSI/gcgdata/geneseg/geneseqn-emb1/NA2002.DAT.*
25: /SIDSI/gcgdata/geneseg/geneseqn-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed.

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1450	100.0	834	23	AAK98735	DNA of a human uri
2	1450	100.0	1288	22	AA158658	Human polynucleoti
3	1450	100.0	1624	24	AA271786	Human 57658 DNA.
4	1450	100.0	2152	22	AA160444	Human polynucleoti
5	1450	100.0	2160	22	AA156521	Human cDNA sequenc
6	1380	95.2	1396	24	AA189762	Human polynucleoti
7	1372.5	94.7	900	22	AAH75355	Human uridine kina
8	1074.5	74.1	753	22	AAH04832	Human cDNA clone (
9	941.5	64.9	1332	22	AAH23801	Human transferase
10	940	64.8	1402	22	AA157850	Human polynucleoti
11	924	63.7	734	24	AB551801	Human mdct cDNA in
12	907.5	62.6	1332	24	AB054470	Human ovorian anti
13	902.5	62.2	1310	21	AA15785	Human prostate can
14	794.5	54.8	1316	22	AA159636	Human polynucleoti
15	738	50.9	822	23	AB110981	Drosophila melanog
16	736.5	50.8	549	25	ABT33606	Anticancer agent t
17	701.5	48.4	715	20	AA215454	Human gene express
18	647	44.6	19125	22	AA136440	Human musculooske
19	647	44.6	19125	25	ABX59428	CDNA encoding nove
20	622.5	42.9	3413	23	AB110980	Drosophila melanog
21	590	40.7	980	23	AA591403	DNA encoding novel
22	578	39.9	1770	21	AAAC4047	Zea mays DNA fragm
23	570.5	39.3	1536	21	AAAC48913	Arabidopsis thalia
24	561.5	38.7	1940	21	AAAC3670	Zea mays DNA fragm
25	545.5	37.6	1847	24	ABX90806	CDNA encoding huma
26	544	37.5	1401	24	AB214138	Arabidopsis thalia
27	542	37.4	1856	23	AB120375	Drosophila melanog
28	542	37.4	2319	23	AB120375	Drosophila melanog
29	542	37.4	3193	23	AB108329	Arabidopsis thalia
30	540.5	37.3	1813	24	ABX59846	Novel human coding
31	540.5	37.3	1959	22	AA545081	CDNA encoding nove
32	529	36.5	1326	21	AAAC50684	Arabidopsis thalia
33	529	36.5	1831	21	AAAC37084	Arabidopsis thalia
34	526.5	36.3	2065	21	AAAC75866	Human ORF1421
35	518	35.7	1503	21	AAAC48458	Arabidopsis thalia
36	507.5	35.0	1500	21	AAAC49434	Arabidopsis thalia
37	507.5	35.0	1703	21	AAAC42628	Arabidopsis thalia
38	490.5	33.8	1838	21	AAAC51009	Arabidopsis thalia
39	485.5	33.5	3881	23	ABV23998	Human prostate exp
40	485.5	33.5	3881	23	ABV24102	Human prostate exp
41	485.5	33.5	3881	23	ABV24281	Human prostate exp
42	485.5	33.5	3881	23	ABV29881	Human prostate exp
43	485.5	33.5	3881	23	ABV29883	Human prostate exp
44	485.5	33.5	3881	23	ABV30091	Human prostate exp
45	473.5	32.7	7248	23	AB108328	Drosophila melanog

ALIGNMENTS

RESULT 1
AAK98735
ID AAK98735 standard; DNA; 834 BP.
XX AAK98735,
XX
XX
XX 02-MAY-2002 (first entry)
XX
XX
XX DNA of a human uridine kinase (UDK).
XX
XX
XX Human; uridine kinase; diagnostic assay; mutation detection; UDK;
XX probab; chromosome localisation study; tissue expression; gene therapy;
XX antibody; vaccine; human ovarian cancer; immunological disorder;
XX human colon carcinoma; immunogen; ds.
XX
XX Homo sapiens.
XX OS

FH Key Location/Qualifiers
 FT CDS 1..780
 FT /tag=a
 FT /partial
 FT /note="No stop codon"
 XX MO200172963-A2.
 XX
 PD 04-OCT-2001.
 XX
 PF 27-MAR-2001; 2001WO-US09663.
 XX
 PR 27-MAR-2000; 2000US-0536647.
 XX
 PA (SMIK) SMITHKLINE BEECHAM CORP.
 XX
 PI Ho YS, Johnson RK;
 XX
 DR WPI: 2001-626259/72.
 DR P-PSDB; AAO14412.
 PT Novel human uridine kinase polypeptides useful for treating cancers,
 PT and to identify agonists and antagonists of the polypeptide useful for
 PT treating conditions associated with uridine kinase imbalance -
 XX
 PS Claim 5; Page 22-23; 31pp; English.
 XX
 CC The invention relates to newly identified human uridine kinase (UDK)
 CC polypeptides and polynucleotides and methods for producing such
 CC polypeptides by recombinant techniques. Also disclosed in the invention
 CC are methods for utilizing uridine kinase polypeptides and polynucleotides
 CC in diagnostic assays. The polynucleotides and polypeptides of the
 CC invention may be used as diagnostic reagents by detecting mutations in an
 CC associated gene. An array of oligonucleotide probes comprising the
 CC uridine kinase polynucleotide sequence or fragments thereof can be
 CC constructed to conduct efficient screening or genetic mutations, for
 CC example. Detection of abnormally decreased or increased levels of
 CC polypeptide or mRNA expression may also be used for diagnosing or
 CC determining susceptibility of a subject to a disease of the invention.
 CC The polynucleotide sequences of the invention can be used for chromosome
 CC localisation studies and tissue expression studies. The polypeptides of
 CC the invention or fragments thereof may be used as immunogens to produce
 CC antibodies. These antibodies may be employed to isolate or identify
 CC clones expressing the polypeptide. The polypeptides and polynucleotides
 CC of the invention can be used as a vaccine or in gene therapy to treat
 CC diseases such as human ovarian cancer, human colon carcinomas, and
 CC immunological disorders. This polynucleotide sequence represents the DNA
 CC of a human uridine kinase of the invention.
 XX
 SO Sequence 834 BP; 194 A; 219 C; 265 G; 156 T; 0 other;
 XX
 Alignment Scores:
 Pred. No.: 9.71e-154 Length: 834
 Score: 1450.00 Matches: 277
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 23 Gaps: 0
 US-09-896-522-2 (1-277) x AAK98735 (1-834)
 QY 1 MetAlaSerAlaGlyGlyGluAspGluSerProAlaProGluAlaAspArgProHis 20
 DB 1 ATGGCTTCGGCGGAGCGGAGACGTCGAGAGCCCGCGGAGCGCGACCGTCGCCGAC 60
 QY 21 GlnArgProPheLeuLeuLeuGlyValSerGlyGlyThrAlaSerGlyLysSerThrValCys 40
 DB 61 CAGCGGCCCTTCCTCATAGGGGTGAGCGCGGCACTGCGACGGGAAGTCCACCGTGTGT 120
 QY 41 GlnLysIleMetGluLeuLeuGlyGlnAsnGluValGlnGlnArgGlnArgLysValVal 60
 DB 121 GAGAAAGATCATGAGATTCTCTGGGACAGAAAGAGGTGAAACAGCGGCGGAGAGGTGCTC 180

QY 61 IleLeuSerGlnAspArgPheTyrLysValIleuThrAlaGluGlnLysAlaLysAlaLeu 80
 DB 181 ATCTTAGCCAGGACAGGTTCTACAGAGTCTGACCGCAGAGCAAGGCCAGGCTTG 240
 QY 81 LysGlyGlnTyrAsnPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThr 100
 DB 241 AAGAGCAAGTACAAATTTTGTACATCCAGATGCTTTGACAAATGATTTGTATGACAGACT 300
 QY 101 LeuLysAsnIleValGluGlyLysThrValGluValProThrTyrAspPheValThrHis 120
 DB 301 CTGAGAGACATCGTGGAGGGCAAAACGTTGAGAGTCCGACCTATGATTTTGTACACAC 360
 QY 121 SerArgLeuProGluThrThrValValTyrProAlaAspValValLeuPheGluGlyIle 140
 DB 361 TCAAGGTACACAGACACAGGTGTTCTACCTCGGACGTGGTCTGTTTGAGGACATC 420
 QY 141 LeuValPheTyrSerGlnGluIleArgAspMetPheHisLeuArgLeuPheValAspThr 160
 DB 421 TTGGTGTTCACAGCCAGAGGATCCGGACATGTTCCACTGGGCTCTTCGTGACAC 480
 QY 161 AspSerAspValArgLeuSerArgValLeuArgAspValArgArgGlyArgAspLeu 180
 DB 481 GACTCCGACGTCAGGCTGTCTCGAAGAGTTCTCCGGAGCGTGGCCGAGGAGGACCTG 540
 QY 181 GluGlnIleLeuThrGlnTyrThrThrPheValLysProAlaPheGluGluPheCysLeu 200
 DB 541 GACGAGATTCTGACGACGATCACACCTCTGTGTAACCGGCTTCGAGAGTTCTGCTG 600
 QY 201 ProThrLysLysTyrAlaAspValIleIleProArgGlyValAspAsnMetValAlaIle 220
 DB 601 CCGAACAAAGAGATGCGCATGATCATCCACGAGAGTGAGCAATATGTTGCCATC 660
 QY 221 AsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysLysTyrPheArg 240
 DB 661 AACCTGCTGTGACGACATCCAGCAATTCGAAATGGTGAACATTCGCAATGGCACCGA 720
 QY 241 GlyGlySerAsnGlyArgSerTyrLysArgThrPheSerGluProGlyLysAspHisProGly 260
 DB 721 GAGGGGTCCAAATGGCGGAGCTTCAAGCGGACCTTTTTCAGCCAGGAGCCACTGGG 780
 QY 261 MetLeuThrSerGlyLysArgSerHisLeuGluSerSerSerArgProHis 277
 DB 781 ATGTGACTCTCTGCAAAACGGTCAATTTGAGATCCAGACGAGACCCCA 831
 RESULT 2
 AAI58658 standard; cDNA; 1288 BP.
 AAI58658;
 AC XX
 DT 22-OCT-2001 (first entry)
 XX
 XX Human polynucleotide SEQ ID NO 861.
 DE XX
 XX Human; noctropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; hemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemofactis;
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia; ss.
 KW XX
 OS Homo sapiens.
 XX
 PN WO200153312-A1.
 XX
 PD 26-JUL-2001.
 PF 26-DEC-2000; 2000WO-US34263.
 XX
 PR 21-JAN-2000; 2000US-0488725.
 PR 25-APR-2000; 2000US-0552317.
 PR 09-JUL-2000; 2000US-0598042.
 PR 19-JUL-2000; 2000US-0620312.

PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
XX
XX (HYSB-) HYSBQ INC.
XX
PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou F, Goodrich R, Drmanac RT;
XX
DR WPI; 2001-442253/47.
DR P-PSDB; AAM39502.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
PS
PS Claim 1; SEQ ID NO 861; 10078bp; English.
XX
XX The invention relates to human nucleic acids (AA157798-AA161369) and
CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC activation/inhibition activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX
XX Sequence 1288 BP; 275 A; 361 C; 418 G; 234 T; 0 other;
SQ
XX
XX Alignment Scores:
Pred. No.: 1,79e-153 Length: 1288
Score: 1450.00 Matches: 277
Percent Similarity: 100.00% Conservat: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 22 Gaps: 0
US-09-896-522-2 (1-277) x AA158658 (1-1288)
QY 1 MetAlaSerAlaGlyGlyGluAspCysGluSerProAlaProGluAlaAspArgProHis 20
DB 95 ATGGCTTCGGCGGAGGCGAAGACTGCGAAGCCCCCGCGGAGGCGCGCGCTCCGCAC 154
QY 21 GlnArgProPheLeuIleGlyValSerGlyGlyThrAlaSerGlyLysSerThrValCys 40
DB 155 CACGGCGCTTCCTGATAGGGGTGAGCGGCGGCACTCCAGCGGGAAGTCGACCGTGTGT 214
QY 41 GlnLysIleMetGluLeuLeuGlyGlnAsnGluValGluGlnArgGlnArgLysValVal 60
DB 215 GAAGAAGTCAATGAGTCTGCTGGACAGAAACAGGTGGAACAGCGGCGACGGAAGGTGCTC 274
QY 61 IleLeuSerGlnAspArgPheTyrLysValLeuThrAlaGluGlnLysAlaLysAlaLeu 80
DB 275 ATCTTAGCCAGACAGACAGAGTCTTCAAGAGTCTTGAACGACAGACAGAGGCGGCTTG 334
QY 81 LysGlyGlnTyrAsnPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThr 100
DB 335 AAAGACAGTACATTTTGAACATCCAGATGCCCTTATATGATTTTGAATGACACAGACT 394
QY 101 LeuLysAsnIleValGluGlyLysThrValGluValProThrTyrAspPheValThrHis 120
DB 395 CTGAAGAACAATCGTGGAGGCAAAACGCTGAGAGTGGCGACTATGATTTTGTGACACAC 454
QY 121 SerArgLeuProGluThrThrValValTyrProAlaAspValValLeuPheGluGlyIle 140

DB 455 TCAGAGTTACAGACAGACACGCGTGTCTACCTCGCGGACGTGGTTCTGTTAGAGGCATC 514
QY 141 LeuValPheTyrSerGlnGluIleArgAspMetPheHisLeuArgLeuPheValAspThr 160
DB 515 TTGGTGTCTACAGCAGACAGATCCCGGACATGTTCCACTCGCGCTTCTGTGGACACC 574
QY 161 AspSerAspValArgLeuSerArgValIleuArgAspValArgArgGlyArgAspLeu 180
DB 575 GACTCCGACGTCAAGGCTGTCTGAGAGATTCTCCGGAGCGTGGCCGAGGAGGACCTG 634
QY 181 GluGlnIleLeuThrGlnTyrThrThrPheValLysProAlaPheGluGluPheCysLeu 200
DB 635 GAGCAGATTCTGACGAGACATCCACCTTCGTGAACCCGCTTCAGAGAGTCTGCTG 664
QY 201 ProThrLysArgTyrAlaAspValIleIleProArgGlyValAlaAsnMetValAlaIle 220
DB 695 CCGACAAAGAAAGTATCCGATGTGATCATCCACGAGAGTGGACAATATGTTGGCCATC 754
QY 221 AsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysLysThrPheIle 240
DB 755 AACCTGATGTGACAGACATCCAGACATTCGAATGTGACATTCGAAATGGCACCA 814
QY 241 GlyLysSerAsnGlyArgSerTyrLysArgThrPheSerGluProGlyAspHisProGly 260
DB 815 GAGGGTCCAAATGGCGGAGCTTCAAGCCGACCTTTCTGAGCCAGGAGGACCACTTGG 874
QY 261 MetLeuThrSerGlyLysArgSerHisLeuGluSerSerSerArgProHis 277
DB 875 ATCTGACCTCTGGCAACGCTCACTTTGAGTCCAGACAGACACCCAC 925
RESULT 3
AAD27186
ID AAD27186 standard; DNA; 1624 BP.
XX
XX AAD27186;
AC
XX
XX 18-APR-2002 (first entry)
DT
XX
XX Human 57658 DNA.
DE
XX
XX Human; uridine kinase-like protein; haematopoietic neoplastic disorder;
KW 57658 protein; leukaemia; haematopoietic disorder; diabetes mellitus;
KW rheumatoid arthritis; psoriasis; dermatitis; osteoporosis; rickets;
KW sarcoma; myocardial infarction; hypertension; atherosclerosis;
KW Alzheimer's disease; Parkinson's disease; Huntington's disease;
KW gene therapy; chromosome mapping; tissue typing; dermatological;
KW cytostatic; osteopathic; cardiac; neuroprotective; nootropic;
KW anticovulant; ds.
XX
XX Homo sapiens.
OS
XX
XX Key
FH CDS
FT 94..927
FT /tag= a
FT /product= "Human 57658 protein"
FT misc_feature
FT 94..924
FT /tag= b
FT /note= "This region is specifically claimed as
FT SEQ ID NO: 3 in claim 1 of the specification"
PN WO200202761-A2.
XX
XX 10-JAN-2002.
PD
XX
XX 28-JUN-2001; 2001WO-US21063.
PF
XX
XX 30-JUN-2000; 2000US-216503P.
PR
XX
XX (MILL-) MILLENNIUM PHARM INC.
PA
XX
XX Glucksmann MA;
PI

DR WPI: 2002-140091/18.
 DR P-PSDB; AAE16592.
 XX New isolated human uridine kinase family polypeptide 57658, useful for
 PT creating hematopoietic neoplastic disorders and disorders of neurons,
 PT heart and blood vessels -
 XX
 PS Claim 1; Fig 1a; 103pp; English.
 XX
 CC The patent discloses human uridine kinase-like polypeptides, designated
 CC 57658 and polynucleotides encoding such proteins. 57658 DNAs and proteins
 CC are useful for developing diagnostic and therapeutic agents for 57658-
 CC mediated or related disorders such as hematopoietic neoplastic disorders
 CC (e.g. leukaemia), hematopoietic disorders (e.g. psoriasis, dermatitis,
 CC diabetes mellitus, rheumatoid arthritis, disorders of bone (e.g. rickets,
 CC osteoporosis), heart (e.g. sarcoma, myocardial infarction), blood vessels
 CC (e.g. hypertension, atherosclerosis) and neurons (e.g. Huntington's
 CC disease, Alzheimer's disease, Parkinson's disease). 57658 sequences or
 CC their antibodies are useful in screening assays, detection assays (e.g.
 CC forensic biology) and predictive medicine (e.g. diagnostic assays,
 CC prognostic assays, and monitoring clinical trials and pharmacogenomics).
 CC They are useful as reagents for diagnosing and treating 57658-mediated
 CC disorders. 57658 DNAs are useful in gene therapy, in chromosome mapping,
 CC to identify an individual from a minute biological sample (tissue typing)
 CC and to aid in forensic identification of the biological sample. The
 CC present sequence is a DNA encoding human 57658 protein.

XX Sequence 1624 BP; 354 A; 427 C; 506 G; 337 T; 0 other;

Alignment Scores:

Pred. No.:	2,486-153	Length:	1624
Score:	1450.00	Matches:	277
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	24	Gaps:	0

US-09-896-522-2 (1-277) x AAD27186 (1-1624)

QY	1	MetAlaSerAlaGlyGlyGluAspCysGluSerProAlaProGluAlaAspArgProHis	20
DB	94	ATGGCTTCGGCGGGGCGGAAGACTGCGAGCCCGCGCGGCGGACCGTCCGCAC	153
QY	21	GlnArgProPheLeuIleGlyValSerGlyGlyThrAlaSerGlyLysSerThrValCys	40
DB	154	CAGCGCCCTTCGATGAGGTGACGGCGGCGGACGCGGAGAGTCAAGCTGGT	213
QY	41	GluLysIleMetGluLeuLeuGlyGlnAsnGluValGluGlnArgGlnArgLysValAla	60
DB	214	GAGAAAGATCATGAGATTGCTCGGACAGAACGAGGTGGAACGCGGACGGAAGTGGTC	273
QY	61	IleLeuSerGlnAspArgPheTyrLysValLeuThrAlaGluGlnLysAlaLeu	80
DB	274	ATCCGAGCCGAGGACAGGTTCTACAAAGTCTCTGACGCGAGGAGGCGCAAGGCCCTTG	333
QY	81	LysGlyGlnTyrAsnPheAspHisProAspAlaPheAspAsnAlaPheMetHisArgThr	100
DB	334	AAAGACAGTAAATTTGACATCCAGATGCTTTGATTAATGATTGATCAGACGAGCT	393
QY	101	LeuLysAsnIleValGluGlyLysThrValGluValProThrTyrAspPheValThrHis	120
DB	394	CTGAAGAACATCGTGGAGGCAAAACGTTGAGGCGGACCTAATGATTTTGACACAC	453
QY	121	SerArgLeuProGluThrThrValValTyrProAlaAspValValLeuPheGluGlyTle	140
DB	454	TCAAGGTATCACAGACACACGTTGTTACCTCGGACGCGTTCGTTTGAGGGGATC	513
QY	141	LeuValPheTyrSerGlnGluIleArgPheMetPheHisLeuArgLeuPheValAspThr	160
DB	514	TTGGTGTCTACGACGAGGAGATCCGGGACATGTTCCACCTGCGCTTTGAGACACC	573
QY	161	AspSerAspValArgLeuSerArgArgValLeuArgAspValArgArgGlyArgAspLeu	180

DB	574	GACTCCGAGCTACGCGTCTCGAAGAGTTCTCCGGGACGTGCGCGGAGGAGGACCTG	633
QY	181	GluGlnIleLeuThrGlnTyrThrPheValLysProAlaPheGluGluPheCysLeu	200
DB	634	GAGCAGATTCCTGACGAGTACACACACCTCGTGAACCCGCTTCAGAGATTCCTCG	693
QY	201	ProThrLysLysTyrAlaAspValIleIleProArgGlyValAspAsnMetValAlaIle	220
DB	694	CCGACAAAGAGTATGCCATGTGATCATCCACGAGAGTGAACAATATGTTGCCATC	753
QY	221	AsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysLysTyrPheArg	240
DB	754	AACCTGATGCTGACGACATCCAGGACATTCGATATGTGACATCTGCAATGACACCGA	813
QY	241	GlyGlySerAsnGlyArgSerTyrLysArgThrPheSerGluProGlyLysPheHisProGly	260
DB	814	GAGGGGTCCATGGGCGGAGCTCAAGCCGACCTTTTCGAGCAGAGGAGCCACCTGGG	873
QY	261	MetLeuThrSerGlyLysArgSerHisIleGluLysSerSerArgProHis	277
DB	874	ATGCTGACCTCTGGCAACGCTACATTTGAGTCCAGACGACCCAC	924

RESULT 4
 ID AAI60444
 AAI60444 standard; cDNA; 2152 BP.
 AC AAI60444;
 DT 22-OCT-2001 (first entry)
 XX
 DE Human polynucleotide SEQ ID NO 4433.
 XX
 KW Human; noctropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; hemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokine; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO20015312-A1.
 XX
 PD 26-JUL-2001.
 XX
 PF 26-DEC-2000; 2000WO-US34263.
 XX
 PR 21-JAN-2000; 2000US-0488725.
 PR 25-APR-2000; 2000US-0552317.
 PR 09-JUL-2000; 2000US-0598042.
 PR 19-JUL-2000; 2000US-0620312.
 PR 03-AUG-2000; 2000US-0653450.
 PR 14-SEP-2000; 2000US-0662191.
 PR 19-OCT-2000; 2000US-0693036.
 PR 29-NOV-2000; 2000US-0727344.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Wehman T, Xu C, Xue AJ, Yang Y, Zhang J;
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
 XX
 DR WPI: 2001-442253/47.
 DR P-PSDB; AAM41288.
 PT Novel nucleic acids and polypeptides, useful for treating disorders
 PT such as central nervous system injuries -
 XX
 PS Claim 1; SEQ ID NO 4433; 10078pp; English.
 CC The invention relates to human nucleic acids (AA157798-AA161369) and
 CC the encoded polypeptides (AAM38642-AAM42213) with noctropic,
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful

Alignment Scores:

Pred. No.:	3,76-153	Length:	2160
Score:	1450.00	Matches:	277
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
		Gaps:	0

US-09-896-522-2 (1-277) x ABL15621 (1-2160)

```

QY 1 MetAlaSerAlaGlyGlyGluAspCysGluSerProAlaProGluAlaAspArgProHis 20
   |||||
Db 95 ATGGCTTCGGCGGAGCGGAAGACTGCGAGAGCCCGCGCGAGGCGGACCGTCCGCAC 154
   |||||
QY 21 GlnArgProPheLeuIleGlyValSerGlyGlyThrAlaSerGlyLysSerThrValCys 40
   |||||
Db 155 CAGCGCGCCCTTCGTATAGGGGTGAGCGCGGCGACTGCGACGGGGAAGTCCAGCCGTGTGT 214
   |||||
QY 41 GluLysIleMetGluLeuLeuGlyGlnAsnGluValGluGlnArgGlnArgValVal 60
   |||||
Db 215 GAGAGATCATGAGATTGCTGGACAGAACAGGTGGAACAGCGGCGACCGGAGGTGCTC 274
   |||||
QY 61 IleLeuSerGlnAspArgPheThrValLeuThrAlaGluGlnLysAlaLysAlaLeu 80
   |||||
Db 275 ATCTGAGCCAGAGACAGAGTTCACAAAGGTCCTGACGCGCAGAGCAGAAAGCCAGCCCTTG 334
   |||||
QY 81 LysGlyGlnIleThrAsnPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThr 100
   |||||
Db 335 AAAGACAGTACAAATTTGACCATTCACATGCTTGTATGATTGATGATGACAGGACT 394
   |||||
QY 101 LeuLysAsnIleValGluGlyLysThrValGluValProThrLysAspPheValThrHis 120
   |||||
Db 395 CTGAAGAACAATCTGTCGAGAGGCAAAACGCTGAGAGTCCGACCTATGATTTGTACACAC 454
   |||||
QY 121 SerArgLeuProGluThrThrValValLysProAlaAspValValLeuPheGluGlyIle 140
   |||||
Db 455 TCAAGGTACACAGAGACACACGCTGCTACCCCTGCGGACGTGTTCTGTTGAGGGCATC 514
   |||||
QY 141 LeuValPheTyrSerGlnGluIleArgAspMetPheHisLeuArgLeuPheValAspThr 160
   |||||
Db 515 TTGGTGTTCTACAGCCAGAGATCCGGGACATGTTCCACCTGCGCTCTTGTGTGACACCC 574
   |||||
QY 161 AsperAspValArgLeuSerArgArgValLeuArgAspValArgArgGlyArgAspLeu 180
   |||||
Db 575 GACTCCGACGTCAGGCTGCTCGAAGACTTCTCGGGACGTCGCCGCGAGGAGGACCTG 634
   |||||
QY 181 GluGlnIleLeuThrGlnTyrThrThrPheValLysProAlaPheGluGlnPheCysLeu 200
   |||||
Db 635 GAGCAGATTCTGACCGCAGTACACCCCTTCGTGAAGCCGCGCTTCGAGAGATTCTGCTG 694
   |||||
QY 201 ProThrLysLysTyrAlaAspValIleIleProArgGlyValaAspAsnMetValAlaIle 220
   |||||
Db 695 CCGACAAAGAGATGCGCGATGTGATCATCCACAGAGGTGAGCAATATGCTGGCCATC 754
   |||||
QY 221 AsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysLysTyrHisArg 240
   |||||
Db 755 AACCGATCGTCGACGACATCCAGACATTTCTGAATGTGACATCTGCAATATGGCACCGA 814
   |||||
QY 241 GlyLysSerAsnGlyArgSerTyrLysArgThrPheSerGluProGlyAspHisProGly 260
   |||||
Db 815 GAGAGGTCTCAATGGCGGAGCTACCAAGCGGACCTTTTCTGAGCAGGGGACCACTCTGG 874
   |||||
QY 261 MetLeuThrSerGlyLysArgSerHisIleGlnLysSerSerArgProHis 277
   |||||
Db 875 ATGCTGACCTCTGGCAACGGTCAATTGAGTTCACAGACAGACCCAC 925
   |||||

```

RESULT 6
ABL89762
ID ABL89762 standard; CDNA; 1396 BP.

XX ABL89762;
XX AC ABL89762;

DT 24-MAY-2002 (first entry)

XX DE Human polynucleotide SEQ ID NO 324.

XX KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
 KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor;
 KW vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
 KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; human; secreted protein; gene; ss.

XX OS Homo sapiens.

XX PN WO200190304-A2.

XX PD 29-NOV-2001.

XX PF 18-MAY-2001; 2001WO-US16450.

XX PR 19-MAY-2000; 2000US-205515P.

XX PA (HUMA-) HUMAN GENOME SCI INC.

XX PI Birse CE, Rosen CA;

XX XX WPI; 2002-122018/16.

XX DR P-PSDB; ABB89353.

XX PT Novel 1405 isolated polypeptides, useful for diagnosis, treatment and
 PT prevention of neural, immune system, muscular, reproductive,
 PT gastrointestinal, pulmonary, cardiovascular, renal and proliferative
 PT disorders -

XX PS Claim 4; SEQ ID NO 324; 2081bp + Sequence Listing; English.

XX CC The invention relates to novel genes (ABL89449-ABL90853) and proteins
 CC (ABB89040-ABB90444) useful for preventing, treating or ameliorating
 CC medical conditions e.g. by protein or gene therapy. The genes are
 CC isolated from a range of human tissues disclosed in the specification.
 CC The nucleic acids, proteins, antibodies and (ant)agonists are useful
 CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
 CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
 CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
 CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
 CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
 CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
 CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;
 CC (d) wound healing; (e) neurological diseases e.g. cerebral ischaemia;
 CC and epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
 CC and parasitic infections.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at http://wipo.int/pub/published_pct_sequences.

XX SQ Sequence 1396 BP; 317 A; 358 C; 425 G; 287 T; 9 other;

Alignment Scores:

Pred. No.:	1,566-145	Length:	1396
Score:	1380.00	Matches:	272
Percent Similarity:	98.19%	Conservative:	0
Best Local Similarity:	98.19%	Mismatches:	5
Query Match:	95.17%	Indels:	2
		Gaps:	0

US-09-896-522-2 (1-277) x ABL89762 (1-1396)

```

QY 1 MetAlaSerAlaGlyGlyGluAspCysGluSerProAlaProGluAlaAspArgProHis 20
   |||||
Db 41 ATGGCTTCGGCGGAGCGGAAGACTGCGAGAGCCCGCGCGAGGCGGACCGTCCGCAC 100
   |||||
QY 21 GlnArgProPheLeuIleGlyValSerGlyGlyThrAlaSerGlyLysSerThrValCys 40
   |||||
Db 101 CAGCGCGCCCTTCGTATAGGGGTGAGCGCGGCGACTGTAAGCGGAAATGCAACCGTGTCT 160
   |||||

```

QY 41 GlnuylleuMetGluLeuGluGlnAsnGluValGluGlnArgGlnArgLysValVal 60
 DB 161 GAGAGATCTATGAGATTGCTGGGACAGACAGAGTGGAAACGGCGACGGAGGTGTC 220
 QY 61 IleuSerGlnAspArgPheTyrLysValIleuThrAgiGlnLysAlaLysAlaIleu 80
 DB 221 ATCTGAGCCAGGACGAGGTTCTTAAGAGTCTTGACGGCAGAGCAGAGGCCAAGCCCTTG 280
 QY 81 LysGlnLysIleuThrAsnPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThr 100
 DB 281 AAGAGACAGTACAAATTTTGAACATCCAGATGCCCTTTGATATATATGATGACAGAGACT 340
 QY 101 LeuLysAsnIleValGluGlnLysThrValGluValProThrTyrAspPheValThrHis 120
 DB 341 CTGAGAGACATCTGAGAGGCGAAACGCTGGAGGTGCCGACTTGTATTTTGACACAC 400
 QY 121 SerArgLeuProGlnThrThrValValTyrProAlaAspValValLeuPheGlnLysIle 140
 DB 401 TCAAGGTTACAGAGACAGGCTGTCTACCTCGGACGCTGTTCTGTTGAGGCGATC 460
 QY 141 LeuValPheTyrSerGlnGlnIleArgAspMetPheHisLeuArgLeuPheValAspThr 160
 DB 461 TTGGTTGTTCTACAGCCGAGAGATCCGAGACATGTTCCACCTGCGCTCTTGAGACACC 520
 QY 161 AspSerAspValArgLeuSerArgArgValIleuArgAspValArgArgGlnArgAspLeu 180
 DB 521 GACTCCGACGTCAGGCTGTCTCGAAGAKTCTCCGGAACGTGGCCGAGGAGGAGACTG 580
 QY 181 GlnGlnIleuLeuThrGlnTyrThrThrPheValLysProAlaPheGlnGlnPheCysLeu 200
 DB 581 GAGAGATTTCTGACGCA-TACACACACTTGTGTGMAAGCCGGCTTCGAGAGTTCTGCTG 639
 QY 201 ProThrLysLysTyrAlaAspValIleIleProArgGlnValAlaAsnMetValAlaIle 220
 DB 640 CCGACAAAGAGATATGCGATGATATCCACAGAGAGTGAACAATATGTTGCCATC 699
 QY 221 AsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlnLysAspIleCysLysThrHisArg 240
 DB 700 AACCTGATCTGTCAGACATCCAGACATTTCTGAATGATGATCTCAATGGCACCGA 759
 QY 241 GlnLysSerAsnGlnArgSerTyrLysArgThrPheSerGlnProGlnAspHisProGln 260
 DB 760 GGAGGGTCCAATGGGCGGAG-TACAAAGCGACCTTTTCTRAAGCGAGGACCACTGGG 818
 QY 261 MetLeuThrSerGlnLysArgSerHisLeuGlnLysSerSerArgProHis 277
 DB 819 ATGCTGACCTCTGGCAACCGTCAATTGGAGTCCAGACAGACCCAC 869
 RESULT 7
 AAH75355 ID AAH75355 standard; cDNA; 900 BP.
 AC AAH75355;
 DT 02-OCT-2001 (first entry)
 DE Human uridine kinase encoding cDNA.
 KW Human; uridine kinase; UK; ss.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT CDS 8..838
 FT /*tag= a
 FT /product= "uridine kinase"
 FT /note= "claimed in claim 1"
 PN CNI287172-A.
 XX 14-MAR-2001.
 PD 07-SEP-1999; 99CN-0118818.

XX 07-SEP-1999; 99CN-0118818.
 PR (UYFU-) UNIV FUDAN.
 PA Yu L, Zhao Y, Zhang H;
 PI WPL; 2001-409529/44.
 DR P-PSDB; AAG64506.
 DR Human uridine kinase and its coding sequence, preparation and
 PT application -
 PS Claim 1; Page 15(Disclosure); 20pp; Chinese.
 CC The invention relates to human uridine kinase (UK).
 XX Sequence 900 BP; 201 A; 237 C; 292 G; 170 T; 0 other;
 SQ
 Alignment Scores:
 Pred. No.: 5,92e-145 Length: 900
 Score: 1372.50 Matches: 270
 Percent Similarity: 98.19% Conservative: 2
 Best Local Similarity: 97.47% Mismatches: 5
 Query Match: 94.66% Indels: 3
 DB: 22 Gaps: 0
 US-09-896-522-2 (1-277) x AAH75355 (1-900)
 QY 1 MetAlaSerAlaGlnLysGlnAspCysGlnSerProAlaProGlnAlaAspArgProHis 20
 DB 8 ATGGCTTCGGCGGAGGCGAAGG-TGCGAGAG-CCGCGCCGAGAGG-AAACCTCCGAC 64
 QY 21 GlnArgProPheLeuIleGlnValSerGlnLysThrAlaSerGlnLysSerThrValCys 40
 DB 65 CAGCGTCCCTCTCTGATAGGGGTGAGCGGCGGCACTGCCAGCGGAAGTGCACCTGTGT 124
 QY 41 GlnLysIleuMetGlnLeuLeuGlnGlnIleValGluGlnArgGlnArgLysValVal 60
 DB 125 GAGAGATCTAGAGTTCGAGGACAGAACAGAGTGAACAGCGGACCGGTAAGTGTGTC 184
 QY 61 IleuSerGlnAspArgPheTyrLysValIleuThrAlaGlnGlnLysAlaLysAlaIleu 80
 DB 185 ATCTGAGCCAGGACAGGTTCTTAAGGTCCTGACCGCAGAGAGAAAGCGCAAGCCTTG 244
 QY 81 LysGlnLysIleuThrAsnPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThr 100
 DB 245 AAGAGACAGTACAAATTTTGAACATCCAGATGCTTGTATGATGATTTGATGACAGAGACT 304
 QY 101 LeuLysAsnIleValGluGlnLysThrValGluValProThrTyrAspPheValThrHis 120
 DB 305 CTGAAAGAACATCTGAGAGGCGCAAAACGTGAGAGTCCGACCTATGATTTTGACACAC 364
 QY 121 SerArgLeuProGlnThrThrValValTyrProAlaAspValValLeuPheGlnLysIle 140
 DB 365 TCAAGGTTACAGAGACAGCGGTGCTTACCTCGCGGACGTGTTCTGTTGAGGCGATC 424
 QY 141 LeuValPheTyrSerGlnGlnIleArgAspMetPheHisLeuArgLeuPheValAspThr 160
 DB 425 TTGGTGTCTACAGCAGAGAGATCCGGACATGTTCCACTGGGCTCTTCGTGACACC 484
 QY 161 AspSerAspValArgLeuSerArgValIleuArgAspValArgArgGlnArgAspLeu 180
 DB 485 GACTCCGAGTCAGGCTGTCTCGAAGAGTTCCTCGGAGCGTGGCGCGGAGGAGACTG 544
 QY 181 GlnGlnIleuThrGlnTyrThrThrPheValLysProAlaPheGlnGlnPheCysLeu 200
 DB 545 GAGCAATTTCTAGCGAGTACACACCTTCGTGAAGCCGCGCTTCAGAGAGTTCTGCTG 604
 QY 201 ProThrLysLysTyrAlaAspValIleIleProArgGlnValAlaAsnMetValAlaIle 220
 DB 605 CCGACAAAGAGATATCCGATGATCATCCACAGAGAGTGAACAATATGTTGCGATC 664

QY 221 AsnIleuValGlnHisIleGlnAspIleuAsnGlyAspIleCysLysTrpHisArg 240
 Db 665 AACCTGATCGTCGACGACATCCAGACATCTGAAATGTGACATCTGCAAAATGGCACC 724
 QY 241 GlyIysSerAsnGlyArgSerTyrIysArgThrPheSerGluProGlyAspHisProGly 260
 Db 725 GGAGGCTCCAACTGGCGCGACGACGACCTTTCTGTGACGACGAGGACCACTCTGG 784
 QY 261 MetLeuThrSerGlyLysArgSerHisIleuGluSerSerArgProHis 277
 Db 785 ATGCTGACCTCTGGCAACGGTCACTTGTGAGTCCAGACAGACCCAC 835
 RESULT 8
 AAH04832
 ID AAH04832 standard; cDNA; 753 BP.
 XX
 AC AAH04832;
 XX
 DT 26-JUN-2001 (first entry)
 XX
 DE Human cDNA clone (5'-primer) SEQ ID NO:1667.
 XX
 KM Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
 XX
 OS Homo sapiens.
 XX
 PN EPI074617-A2.
 PD 07-FEB-2001.
 XX
 PF 28-JUL-2000; 2000EP-0116126.
 XX
 PR 29-JUL-1999; 99JP-0248036.
 PR 27-AUG-1999; 99JP-0300253.
 PR 11-JAN-2000; 2000JP-0118776.
 PR 02-MAY-2000; 2000JP-0183767.
 PR 09-JUN-2000; 2000JP-0241899.
 XX
 PA (HELI-) HELIX RES INST.
 XX
 PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX
 DR WPI; 2001-318749/34.
 XX
 PT Primer sets for synthesizing polynucleotides, particularly the 5602
 PT full-length cDNAs defined in the specification, and for the detection
 PT and/or diagnosis of the abnormality of the proteins encoded by the
 PT full-length cDNAs -
 XX
 PS Claim 1; SEQ ID 1667; 2537pp + CD ROM; English.
 XX
 CC The present invention describes primer sets for synthesizing 5602
 CC full-length cDNAs defined in the specification. Where a primer set
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
 CC to the complementary strand of a polynucleotide which comprises one of
 CC the 5602 nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination
 CC of the 5'-end sequence/3'-end sequence is selected from those defined in
 CC the specification. The primer sets can be used in antisense therapy and
 CC in gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to
 CC AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632
 CC represent oligonucleotides, all of which are used in the exemplification

CC of the present invention.
 XX
 SQ Sequence 753 BP; 158 A; 196 C; 261 G; 135 T; 3 other:
 Alignment Scores:
 Pred. No.: 1,83e-111 Length: 753
 Score: 1074.50 Matches: 213
 Percent Similarity: 97.26% Conservative: 0
 Best Local Similarity: 97.26% Mismatches: 3
 Query Match: 74.10% Indels: 3
 DB: 22 Gaps: 2
 US-09-896-522-2 (1-277) x AAH04832 (1-753)
 QY 1 MetAlaSerAlaGlyGlyGluAspCysGluSerProAlaProGluAlaAspArgProHis 20
 Db 95 ATGGCTTCGGCGGAGCGCAAGACTGTGAGAGCCCGCGCGAGGCCGCGCTCGCAC 154
 QY 21 GlnArgProPheLeuIleGlyValSerGlyGlyThrAlaSerGlyLysSerThrValCys 40
 Db 155 CAGCGGCCCTTCTGTGATAGGGGTGAGCGGCGGACACTGCCAGCGGGAAGTGACCGTGT 214
 QY 41 GluLysIleMetGluLeuLeuGlyGlnAsnGluValGluGlnArgGlnArgLysVal 60
 Db 215 GAGAGATCATGAGATGCTGCGGACAGAACGAGTGGAACCGCGACGCGAGCTGCTC 274
 QY 61 IleLeuSerGlnAspArgPheTyrLysValIleThrAlaGluGlnLysAlaLysValLeu 80
 Db 275 ATCCTGAGCCAGACAGGCTTCTCAAGGCTCTACGCGACAGACGAAAGCCAGGCTTG 334
 QY 81 LysGlyGlnTyrAsnPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThr 100
 Db 335 AAGGACAGTACATATTTTGACCATCCAGATCCCTTGATATATGATTTGACACAGACT 394
 QY 101 LeuLysAsnIleValGluGlyLysThrValGluValProThrTyrAspPheValThrHis 120
 Db 395 CTGAGAGACATCGTGGAGGGGCAAAACGTGGAGAGTCCGACCTATGATTTTGACACAC 454
 QY 121 SerArgLeuProGluThrThrValValTyrProAlaAspValIleuPheGluGlyIle 140
 Db 455 TCAAGGTACCAAGACACCGGTGTCTACCTCGGACGCTGTTCTGTTGGGGCATC 514
 QY 141 LeuValPheTyrSerGlnGluIleArgAspMetPheHisLeuArgLeuPheValAspThr 160
 Db 515 TTGGTGTCTACAGCCAGACGATCCGGACATTTCCACTGCGCTCTTCGTGGACACC 574
 QY 161 AspSerAspValArgLeuSerArgArgValLeuArgAspValArgArgLysArgLeu 180
 Db 575 GACTCCGACCTCAGGCTGTCTCCAAAGAGTTCTCCGGGACGTGCGGAGGAGGACCTG 634
 QY 181 GluGlnIleLeuThrGlnTyrThrThrPheValLysProAlaPheGluGluPheCysLeu 200
 Db 635 GAGCAGATTGTGACGCGATACACCATCTTCGTAAACCCGGCTTCAGAGAGTTCC---TTG 691
 QY 201 ProThr-----LysLysTyrAlaAspValIleIleProArgGlyValAspAsnMet 217
 Db 692 CTTTGCACAAAGAAAGTGTCCGATGTGATTCACAGAGAGTGGCAATATAG 748
 RESULT 9
 AAH23801
 ID AAH23801 standard; cDNA; 1322 BP.
 XX
 AC AAH23801;
 XX
 DT 31-JUL-2001 (first entry)
 XX
 DE Human transferrase HTFS-1 cDNA, SEQ ID NO:43.
 XX
 KM Human transferrase; HTFS; agonist; antagonist; cellular signalling;
 KM proliferation; cell proliferative disorder; immune disorder;
 KM atherosclerosis; hepatitis; psoriasis; cancer; tumour;
 KM inflammation; AIDS; Addison's disease; allergy; asthma; anaemia;
 KM cirrhosis; Crohn's disease; atopic dermatitis; diabetes mellitus;

KM multiple sclerosis; rheumatoid arthritis; pancreatitis;
 KM systemic lupus erythematosus; thrombocytopenia; ulcerative colitis;
 KM haemodialysis; extracorporeal circulation; trauma; transgenic animal;
 KM gene therapy; drug screening; ss.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 250..1025
 FT /tag= a
 FT /product= "HTFS-1"
 FT /function= "Transferrase"
 XX
 PN MO200132888-A2.
 XX
 PD 10-MAY-2001.
 XX
 PF 02-NOV-2000; 2000MO-US30485.
 XX
 PR 04-NOV-1999; 99US-0163595.
 XX
 PA (INCY-) INCYTE GENOMICS INC.
 XX
 PI Tang YT, Yue H, Hillman JL, Lal P, Bandman O, Patterson C;
 PI Shih LT, Azimzai Y, Lu DM, Baughn MR;
 XX
 DR WPI: 2001-328796/34.
 DR P-PSDB; AAB73494.
 XX
 PT Human transferrase polypeptides and polynucleotides useful for
 PT diagnosis, prevention and treatment of cell proliferative and immune
 PT system disorders and for identifying agonists and antagonists -
 XX
 PS Claim 5; Page 135-136; 157pp; English.
 XX
 CC Sequences AAB73494-AAB73535 represent novel human transferrase proteins
 CC HTFS-1 to HTFS-42, and sequences AAH23801-AAH23842 represent cDNAs
 CC encoding them. The proteins play important roles in the regulation of
 CC cellular signalling and proliferation. The HTFS proteins are useful for
 CC screening compounds for their effectiveness as agonists or antagonists of
 CC transferrase activity, or for compounds that specifically bind to an HTFS
 CC protein or which modulates the activity of an HTFS protein.
 CC Pharmaceutical compositions comprising an HTFS protein, HTFS
 CC agonist or antagonist, or genetic construct encoding an HTFS
 CC protein are useful for treating a disease or condition associated
 CC with decreased or increased expression of functional HTFS. Disorders
 CC which may be treated using such compositions include cell proliferative
 CC disorders and immune disorders. For example, diseases which may be
 CC treated include atherosclerosis, hepatitis, psoriasis, cancers (including
 CC breast, bladder, bone marrow, brain and uterus cancer), inflammation,
 CC AIDS, Addison's disease, allergies, asthma, anaemia, cirrhosis, Crohn's
 CC disease, atopic dermatitis, diabetes mellitus, multiple sclerosis,
 CC rheumatoid arthritis, pancreatitis, systemic lupus erythematosus,
 CC thrombocytopenia, and ulcerative colitis. They may also be used to treat
 CC complications of cancer, haemodialysis, extracorporeal circulation,
 CC trauma and haematopoietic cancer, including lymphoma, leukemia and
 CC myeloma. Polynucleotides encoding HTFS proteins are useful for creating
 CC transgenic animals to model human diseases, for diagnostic purposes and
 CC to generate hybridisation probes useful in mapping the naturally
 CC occurring genomic sequences. HTFS, and its catalytic or immunogenic
 CC fragments are useful for screening libraries of compounds in a variety of
 CC drug screening techniques. Antibodies which specifically bind HTFS may be
 CC used for the diagnosis of disorders associated with the expression of
 CC HTFS, or in assays to monitor patients being treated with HTFS or
 CC agonists, antagonists or inhibitors of HTFS. The present sequence
 CC represents an HTFS protein-encoding cDNA of the invention.
 XX
 SO Sequence 1322 BP; 339 A; 359 C; 358 G; 266 T; 0 other;
 Alignment Scores:
 Pred. No.: 4e-96 Length: 1332
 Score: 941.50 Matches: 192
 Percent Similarity: 78.87% Conservative: 32

Best Local Similarity: 67.61% Mismatches: 37
 Query Match: 64.93% Indels: 23
 DB: 22 Gaps: 5
 US-09-896-522-2 (1-277) x AAH23801 (1-1322)
 QY 3 SerAlaGlyGlyGluAspCysGluSerProAlaProGluAlaAspArgPro----- 19
 DB 223 GCACGGGAGAGAGGGGGCGGCGAACCAACGCGGAGACGACGACGACCTCGCAAC 282
 QY 20 HsGlnArg-----ProPheLeuIleGlyValSerGlyGlyThraLaser 34
 DB 283 CACACAGACGCCACACGCGCGCAGGCTCTTCTTATAGCGGTACGCGGGAACAGCTAC 342
 QY 35 GlyLysSerThraValCysGluLysIleMetGluLeuLeuGlyGlnAsnGluValGluGln 54
 DB 343 GGCAGGTCTTCGTGTGTCTAAGATCGGACGCTCTGGGACAGATAGAGGTAGCTAT 402
 QY 55 ArgGlnArgLysValValIleLeuSerGlnAspArgPheTyrLysValLeuThraLaglu 74
 DB 403 CGCCAGACAGAGGTGTCATCTCGACCGAGATAGCTTACCGTCTTACCTCGAG 462
 QY 75 GlnLysAlaLysAlaLeuLysGlyGlnTyrAsnPheAspHisProAspAlaPheAspAsn 94
 DB 463 CAGAAAGCCAAAGCCCTGAAGGGCCAGTCACTTGAACACCCGGATGCTTGAACAT 522
 QY 95 AspleuMechIaArgThrLeuLysAsnIleValGluGlyLysThrValGluValProThr 114
 DB 523 GAACCTATCTCAAAACACTCAAAAGAAATCACTGAAGGAGAAACATCCAGATCCCGTG 582
 QY 115 TyrAspPheValThrHisSerArgLeuProGluThrThraValValTyrProAlaAspVal 134
 DB 583 TATGACTTGTCTCCATTCCTCCGGAAGAGAGACAGTACTGTCTATCCCGACGTG 642
 QY 135 ValLeuPheGluGlyIleLeuValPheTyrSerGlnGluIleArgAspMetPheHisLeu 154
 DB 643 GGGCTTGTGAAGAGGATCTGGCGCTTCTACTCCAGAGAGTACGAGACCTGTCCAGATG 702
 QY 155 ArgLeuPheValAspThrAspSerAspValArgLeuSerArgArgValLeuArgAspVal 174
 DB 703 AAGCTTTTGTGGATACATGACATCGGACCGGCTCTCGAGAGATATTAAAGGACATC 762
 QY 175 ---ArgArgGlyArgAspLeuGluGlnIleLeuThrGlnTyrThrPheValLysPro 193
 DB 763 AGCGAGAGGAGGAGGATTTGACGAGATTTTATCTCACTATTCAGTTCGTCAGAGCT 822
 QY 194 AlaPheGluGluPheCysLeuProThrLysLysTyrAlaAspValIleLeuProArgGly 213
 DB 823 GCTTTGAGAGATTCGCTTGCACAAAGAAATGCTGATGATATCCCTAGAGCT 882
 QY 214 ValAspAsnMetValAlaIleAsnLeuIleValGlnHisIleGlnAspIleLeuAsnGly 233
 DB 883 GCAGATAACTGCTGCGCTCAACCTCACTCGAGCAGCATCAGAGACATCTGTAAATGA 942
 QY 234 AspIleCysLysTyrPheHisArgGlyLysSerAsnGlyArgSerTyrLysArgThrPheSer 253
 DB 943 -----GGGCCCCCAAAAGGAGAC----- 963
 QY 254 GluProGlyAspHisProGlyMetLeuThrSerGlyLysArgSerHisLeuGluSerSer 273
 DB 964 ---AATGGTGTCTCAAGCGCTACACCCCTTACCGAAGAGGACAGGCATCGAGTCCAC 1020
 QY 274 SerArgProHis 277
 DB 1021 AGCAGGCCCAT 1032
 RESULT 10
 AA157850 standard; cDNA; 1402 BP.
 ID AA157850
 AC AA157850;
 XX
 XX
 DT 22-OCT-2001 (first entry)

XX DE Human polynucleotide SEQ ID NO 53.
 XX XX
 KW Human, nootropic; immunosuppressant; cyrostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia; ss.
 XX OS Homo sapiens.
 XX PN WO200153312-A1.
 XX PD 26-JUL-2001.
 XX PF 26-DEC-2000; 2000WO-US34263.
 XX PR 21-JAN-2000; 2000US-0488725.
 XX PR 25-APR-2000; 2000US-0552317.
 XX PR 09-JUL-2000; 2000US-0598042.
 XX PR 19-JUL-2000; 2000US-0620312.
 XX PR 03-AUG-2000; 2000US-0653450.
 XX PR 14-SEP-2000; 2000US-0662191.
 XX PR 19-OCT-2000; 2000US-0693036.
 XX PR 29-NOV-2000; 2000US-0727344.
 XX PA (HSE-) HYSEQ INC.
 XX FI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D,
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AD, Yang Y, Zhang J,
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
 XX DR WPI: 2001-4442253/47.
 DR P-PSDB; AAM38694.
 PT Novel nucleic acids and polypeptides, useful for treating disorders
 PT such as central nervous system injuries -
 XX PS Claim 1; SEQ ID NO 53; 10078pp; English.
 XX CC The invention relates to human nucleic acids (AA157798-AA161369) and
 CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
 CC immunosuppressant and cyrostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, Leukemias and
 CC C.N.S disorders.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification.
 XX SQ Sequence 1402 BP; 379 A; 370 C; 371 G; 282 T; 0 other;
 Alignment Scores:
 Pred. No.: 6.42e-96 Length: 1402
 Score: 940.00 Matches: 194
 Percent Similarity: 74.66% Conservative: 27
 Best Local Similarity: 65.54% Mismatches: 41
 Query Match: 64.83% Indels: 34
 DB: 22 Gaps: 5
 US-09-896-522-2 (1-277) x AA157850 (1-1402)
 QY 2 A1aser1ag1y1g1y1g1u-----7
 DB 227 GCGGGCGCGCGCGCGCGTGCCTTCGACAGGACGCGGAGAGGGCGCGCGA 286

QY 8 -----AspCysGluSerProAlaProGluAlaAspArgProHis---GlnArg 22
 DB 287 ACCATGCGCGGGACACGACGACGACCCCTGCAGAACACACGACGACGCGCGGAG 346
 QY 23 ProPheLeuIleGlyValSerGlyGlyThrAlaSerGlyLysSerThrValCysGlyLys 42
 DB 347 CCTTCCTTATAGCGCGACGCGGGGAAACAGTACGCGGCAAGCTTCCTCGTGTGCTAAG 406
 QY 43 IleMetGluLeuLeuGlyGlnIleGlnValGluGlnAlaGlnIleArgLysValValIleLeu 62
 DB 407 ATCGTAGCGCTCCGGGAGAAATGAGTGCATTCGCCAGAAAGAGGTGTCATCTCG 466
 QY 63 SerGlnAspArgPheThrValLeuThrAlaGluGlnLysAlaLysAlaLeuLysGly 82
 DB 467 AGCCAGGATAGCTTCTACCGCTGCTCTTACCTCGGACAGAAAGCCCAAGCCCTGAAAGGC 526
 QY 83 GlnTyrAsnPheAspHisProAspAlaPheAspAsnAspIleuMetHisArgThrLeuLys 102
 DB 527 CAGTTCACCTTGACCAACCGGATGCTTGAACAATGAACTCATTCACAAACACTCAAA 586
 QY 103 AsnIleValGluGlyLysThrValGluValProThrTyrAspPheValThrHisSerArg 122
 DB 587 GAAATCATGTAAGGAAACAGTCCAGATCCCGGTATGACTTGTCTCCACTCCCG 646
 QY 123 LeuProGluThrThrValValTyrProAlaAspValValLeuPheGluGlyIleLeuVal 142
 DB 647 AAGGAGAGACATTAAGTCTGCTATCCCGACAGCTGTGCTCTTTAAGGAGATCTGCGC 706
 QY 143 PheTyrSerGlnGluIleArgAspMetPheHisLeuArgLeuPheValAspThrAspSer 162
 DB 707 TTCTACTCCGAGAGGATGACAGACCTGTTCCAGATGAACTTTTGTGTGATACAGATGCG 766
 QY 163 AspValArgLeuSerArgArgValIleuArgAspVal---ArgArgGlyArgAspLeuGlu 181
 DB 767 GACACCGGCTCTCAGCAGAGATTAAGGACATGACGAGAGAGGAGGATCTTGAG 826
 QY 182 GlnIleLeuThrGlnTyrThrThrPheValLysProAlaPheGluGluPheCysLeuPro 201
 DB 827 CAGATTTCATCTCAGTACATTAAGTCTGCTGCTCAAGCTGCTTTAGAGAAATCTCTGCCA 886
 QY 202 ThrLysLysTyrAlaAspValIleleProArgGlyValAspAsnMetValAlaIleAsn 221
 DB 887 ACAAGAAGATGCTGATGATGATCCATCCAGAGGTGACGATATCTGGTGGCCATCAAC 946
 QY 222 LeuIleValGlnHisIleGlnAspIleLeuAspGlyAspIleCysLysTrpHisArgGly 241
 DB 947 CTCATGTCGACGACATCCAGGACATCTGAATGA-----982
 QY 242 GlySerAsnGlyArgSerTyrLysArgThrPheSerGluProGluLysAspHisProGlyMet 261
 DB 983 -----GGGCGCTCCAAACGACGAC-----AATGGCTGTCTCAACGCGCTAC 1024
 QY 262 LeuThrSerGlyLysArgSerHisLeuGluSerSerSerArgProHis 277
 DB 1025 ACCCTTACAGCAAGACGACGATCGAGTCCAGACAGGCGCGCAT 1072
 RESULT 11
 ID ABS51801 standard; cDNA; 734 BP.
 XX AC ABS51801;
 XX DT 21-OCT-2002 (first entry)
 XX DE Human mdtl cDNA Incyte ID No: LI:235557.12:2001JAN12.
 XX KW Human; molecule for disease detection and treatment; MDT; cancer;
 KW cell proliferative disorder; arteriosclerosis; cirrhosis; asthma;
 KW autoimmune disorder; inflammatory disorders; Crohn's disease;
 KW multiple sclerosis; cyrostatic; antiretroviral; antiinflammatory;
 KW hepatotropic; immunosuppressive; antiaesthetic; gene therapy; gene; ss.

OS Homo sapiens.
 XX PN W0200255738-A2.
 XX PD 18-JUL-2002.
 XX PF 09-JAN-2002; 2002MO-US01008.
 XX PR 12-JAN-2001; 2001US-261622P.
 XX PR 16-JAN-2001; 2001US-261865P.
 XX PR 17-JAN-2001; 2001US-262208P.
 XX PR 17-JAN-2001; 2001US-262209P.
 XX PR 17-JAN-2001; 2001US-262326P.
 XX PR 19-JAN-2001; 2001US-263063P.
 XX PR 19-JAN-2001; 2001US-263065P.
 XX PR 19-JAN-2001; 2001US-263329P.
 XX PA (INCYTE GENOMICS INC.
 XX PI Panzer SR, Lincoln SE, Altus CM, Dufour GE, Hillman JL, Jones AL,
 PI Dam TC, Liu TP, Harris B, Flores V, Dafo A, Marwaha R, Chen AJ,
 PI Chang SC, Gerebin EH, Peralta CH, David MH, Lewis SA;
 XX DR WPI; 2002-590679/63.
 XX DR P-PSDB; ABG70328.
 XX PT New disease detection and treatment molecule (MDPT) polynucleotides and
 PT polypeptides, useful in diagnosing, studying, preventing or treating
 PT diseases associated with MDDR expression, e.g. autoimmune or
 PT inflammatory disorders -
 XX Claim 1; Page 98; 129pp; English.
 XX PS The present invention relates to the isolation of novel human
 XX CC molecules for disease detection and treatment (MDPT), and the
 CC polynucleotide sequences (mdpt) encoding them. The MDDR polypeptides
 CC may be used to screen for molecules that bind to, or are bound by the
 CC encoded polypeptides, and to develop a transcript image of a tissue or
 CC cell type. Probes comprising at least 20 nucleotides of the mdpt
 CC polynucleotide may be used to assess the toxicity of a test compound.
 CC The MDDR polypeptides and mdpt polynucleotides are useful in the
 CC diagnosis, study, prevention and treatment of diseases associated with
 CC the expression of molecules for disease detection and treatment. Such
 CC disorders include cell proliferative disorders (e.g. arteriosclerosis,
 CC cirrhosis, or cancers), and autoimmune/inflammatory disorders
 CC (e.g. asthma, Crohn's disease, or multiple sclerosis). The mdpt
 CC polynucleotide may also be used as molecule markers, in microarrays,
 CC and in somatic or germline gene therapy. ABS51779-ABS51814 encode
 CC the MDDR proteins of the invention.
 XX SQ Sequence 734 BP; 166 A; 181 C; 226 G; 154 T; 7 other;
 Alignment Scores:
 Pred. No.: 1,65e-94 Length: 734
 Score: 924.00 Matches: 183
 Percent Similarity: 87.98% Conservative: 0
 Best Local Similarity: 87.98% Mismatches: 4
 Query Match: 63.72% Indels: 22
 DB: 24 Gaps: 1
 US-09-896-522-2 (1-277) x ABS51801 (1-734)
 QY 90 AspAlaPheAspAsnAspLeuMetHisArgThrLeuLeuAsnIleValGluGlyLeuThr 109
 DB 14 GATGCTTTGATATATGATTGATGACAGACTCTGAAAGAACATCGTGGAGGCGCAAAACG 73
 QY 110 ValGluValProThrTyraAspPheValThrHisSerArgLeuProGluThrThrValVal 129
 DB 74 GTGGAGGTGGCGGACCTATGATTTTGTGACACACTCAAGGTTACAGAGACACAGGTGGTC 133
 QY 130 TyrProAlaAspValValLeuPheGluGlyIleLeuValPheTyrSerGlnGluIleArg 149
 DB 134 TACCTGCGGAGCGTGGTTCTGTTTGAGGGGATCTTGCTGTCTTACAGCCAGGATCCCG 193

QY 150 AspMetPheHisLeuArgPhePheValAspThrAspSerAspValArgLeuSerArgArg 169
 DB 194 GACATGTTCCACCTGGCGCTCTTCGTGGACCCGACTCCGAGCTCGAGCTCTCGAAG 253
 QY 170 ValLeuArgAspValAlaArgArgGlyArgAspLeuGluGlnIleLeuThrGlnTyrThrThr 189
 DB 254 GTTCTCCGGGAGCGTGGCGCGGAGGAGGACCTGGAGACATTTCTGACGCAATACACACC 313
 QY 190 PheValIysProAlaPheGluGluPheCysLeuProThr----- 202
 DB 314 TTCGTGAAGCCGCGCTTCAGAGAGTTCTCGTCCGCA-GCAGAGCATCTGACAGGGAAT 372
 QY 203 -----LysTyrTyrAlaAspVal 208
 DB 373 GAGAGTCACGATTGAGCCCAATAGTGGTTGATGAGGGAACAAAGAAAGTATGCGCATGTG 432
 QY 209 IleIleProArgGlyValAspAsnMetValAlaIleAsnLeuIleValGlnHisIleGln 228
 DB 433 ATCATCCCAAGAGAGTGAACATATGTTGCCATCACTGATCTGACGACATCCAG 492
 QY 229 AspIleLeuAsnGlyAspIleCysIlySTyrPHisArgGlyGlySerAsnGlyArgSerTyr 248
 DB 493 GACATTCGAATGTGACATCTGCAAAATGAGCACACCGAGAGGCTCCAAATGGCGGAGCTAC 552
 QY 249 LysArgThrPheSerGluProGlyAspHisProGlyMetLeuThrSerGlyLysArgSer 268
 DB 553 AAGCGGACNTTTTCTGAGCGAGGAGCCACCTCGGATGCTGACNTCTGCAAAACGGTAA 612
 QY 269 HisLeuGluSerSerSerArgPro 276
 DB 613 CATTTGAGNCCAGCNCCTCG 636
 RESULT 12
 ABQ54470
 ID ABQ54470 standard; cDNA; 1322 BP.
 XX ABQ54470;
 AC 22-AUG-2002 (first entry)
 XX Human ovarian antigen HOELP29 cDNA, SEQ ID NO:350.
 XX Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;
 KW ovarian cancer; breast cancer; tumour; reproductive system disorder;
 KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
 KW PCOS; ovarian cyst; dysmenorrhoea; endocrine disorder; infection;
 KW inflammatory condition; immune disorder; blood disorder;
 KW cardiovascular disorder; respiratory disorder; neurological disorder;
 KW gastrointestinal disorder; urinary system disorder; drug screening;
 KW gene therapy; chromosome mapping; forensic analysis;
 KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;
 KW anti-inflammatory; gynaecological; reproductive; chromosome 1p32;
 KW gene; ss.
 XX Homo sapiens.
 XX W0200200677-A1.
 XX 03-JAN-2002.
 XX 07-UTN-2001; 2001MO-US18569.
 XX 07-UTN-2000; 2000US-209467P.
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX Birse CE, Rosen CA;
 XX WPI; 2002-147878/19.
 XX P-PSDB; ABP41393.
 PT Isolated nucleic acid molecules encoding novel ovarian polypeptides.

PT useful in the prevention, treatment and diagnosis of cancer (e.g.
PT ovarian cancer), immune disorders, cardiovascular disorders and
PT neurological diseases -

PS Claim 1; SEQ ID No 350; 2922bp; English.

XX
XX
XX The invention relates to 2175 novel human ovarian antigens (ABP41054-
CC ABP43228) and to cDNAs encoding them (AB054131-AB056305), and also
CC encompasses polypeptides 90% identical and polynucleotides 95% identical
CC to the sequences of the invention. The invention additionally relates to
CC recombinant vectors and host cells comprising human ovarian antigen
CC polynucleotides, antibodies against human ovarian antigens, and the use
CC of ovarian antigen polynucleotides and polypeptides in diagnosing,
CC treating, prognosing or preventing various ovarian and/or breast-related
CC disorders. Such conditions include ovarian cancer and breast cancer, and
CC metastatic tumours of ovarian or breast origin, reproductive system
CC disorders (e.g., infertility, disorders of pregnancy, anovulation,
CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine
CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic
CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
CC vaginitis), immune disorders (e.g., congenital and acquired
CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
CC blood-related disorders (e.g., anaemia), cardiovascular disorders,
CC respiratory disorders, neurological disorders, gastrointestinal disorders
CC and urinary system disorders. Ovarian antigen polypeptides and
CC polynucleotides may also be used in screening for compounds which
CC modulate ovarian antigen expression or activity. The polynucleotides may
CC further be used for gene therapy, chromosome mapping, in the
CC identification of individuals and in forensic analysis, and the
CC polypeptides may be used as food additives or to prepare antibodies
CC useful in disease diagnosis, drug targeting and phenotyping. The present
CC sequence represents cDNA encoding a human ovarian antigen of the
CC invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX
SQ Sequence 1322 BP; 351 A; 351 C; 352 G; 264 T; 4 other;

Alignment Scores:

Pred. No.: 2.73e-92 Length: 1322
Score: 907.50 Matches: 190
Percent Similarity: 78.17% Conservative: 32
Best Local Similarity: 66.90% Mismatches: 39
Query Match: 62.59% Indels: 24
DB: Gaps: 5

US-09-896-522-2 (1-277) x AB054470 (1-1322)

QY 3 SerAlaGlyGlyGluAseProCysGluSerProAlaProGluAlaAspArgPro----- 19
DB 205 GCACCGGGAGAGAGGGCGCGCGCAACCTATGCGCGGACGACGACGACCTTCGACAGAC 264
QY 20 HisGlnArg-----ProPheLeuIleGlyValSerGlyGlyThrAlaSer 34
DB 265 CACCAGAGCCCAACGCGCGCGACGCCCTTCTTATAGCGCTGAC-GGGGGAACGCTAGC 323
QY 35 GlyIysSerThrValCysGluValIleMetGluLeuLeuGlyGlnAenGluValGluGln 54
DB 324 GGCAGAGCTTCGCGGTGCTGCTTAAGATCGTCAGCTCGGGGAGAAAGAGTGAAGTAT 383
QY 55 ArgGlnArgIysValAlaIleLeuSerGlnAspArgPheTyrIysValIleuThrAlaGlu 74
DB 384 CGCAGAGAGCGAGTGTCTATCTGACGACGATAGCTTTCACCTGCTCTTACCTCGAG 443
QY 75 GlnIysAlaIysAlaLeuIysGlyGlnTyrAsnAspHisProAspAlaPheAspAsn 94
DB 444 CAGAAAGCCAAAGCCCTGAAGGSCCACTTTCACCCGATGCTTGAACAT 503
QY 95 AspLeuMetHisArgThrIleuIysAsnIleValGluGlySerThrValGluValProThr 114
DB 504 GAACCTATTCTCAAAACACTCAAGAAATCACTGAAGGAAACAGATCCAGATCCCGTG 563

QY 115 TyrAspPheValThrHisSerArgLeuProGluThrThrValValTyrProAlaAspVal 134
DB 564 TATGACTTTTCTCCCATTTCCCGAAGAGAGACAGTACTCTCTATCCCGAGACGTG 623
QY 135 ValLeuPheGluGlyIleLeuValPheTyrSerGlnGluIleArgAspMetPheHisLeu 154
DB 624 GTGCTCTTGAAGGATCCCTGGCCTTCTACTCCAGAGAGTACGAGACCTGTTCCAGATG 683
QY 155 ArgLeuPheValAspThrAspSerAspValArgLeuSerArgArgValLeuAlaAspVal 174
DB 684 AAGCTTTTGTGATACAGATGCGGACACCCGCTCTCAGCGAGAGATTAAAGGACATC 743
QY 175 ---ArgArgGlyArgAspLeuGluGlnIleLeuThrGlnTyrThrThrPheValIysPro 193
DB 744 AGCAGAGAGAGCGAGATCTTGACGACATTTTATCTCAGTACATTACGTTCTGCAAGCT 803
QY 194 AlaPheGluGluPheCysLeuProThrIysIysTyrAlaAspValIleIleProArgGly 213
DB 804 GCTTTGAGGAATTCTGCTGCGCAACAGAAAGATAGTGTATGATCCCTAGAGGT 863
QY 214 ValAspAspMetValAlaIleAenIleValGlnHisIleGlnAspIleLeuAenGly 223
DB 864 GCAGATATATCTGGTGGCCATCAACCTCATCTGTCACACATCCAGACATCTGAATGGA 923
QY 234 AspIleCysIysThrPheAspArgIysSerAsnGlyArgSerTyrIysArgThrPheSer 253
DB 924 -----GGGCTCTCAAAAGGCGAGAC----- 944
QY 254 GluProGlyAspHisProGlyMetLeuThrSerGlyIysArgSerHisLeuGluSer 273
DB 945 ---AATGGCTGTCTCAACGGCTACACCCCTTACGCAAGAGGAGGACATCGAGTCCAGC 1001
QY 274 SerArgProHis 277
DB 1002 AGCAGCGCGCAT 1013
RESULT 13
ID AAF15785 standard; cDNA; 1310 BP.
XX AAF15785;
AC 13-MAR-2001 (first entry)
DT 13-MAR-2001 (first entry)
XX
DE Human prostate cancer antigen nucleotide sequence SEQ ID NO:220.
XX
XX Human; prostate cancer; prostate cancer antigen; detection; diagnosis;
XX neuroprotective; cytostatic; cardiocactive; immunomodulatory; muscular;
XX vulnery; gastrointestinal; nephrotoxic; antiinfective; gynaecological;
XX antibacterial; gene therapy; neural; immune; reproductive; renal;
XX gastrointestinal; pulmonary; cardiovascular; proliferative disorder;
XX wound; infectious disease; se.
XX
OS Homo sapiens.
PN WC200055174-A1.
XX
XX 21-SEP-2000.
XX
XX 08-MAR-2000; 2000MO-US05988.
XX
XX 12-MAR-1999; 99US-0124270.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX (ROSE/) ROSEN C A.
XX
XX Rosen CA, Ruben SM;
XX
XX WPI; 2000-587513/55.
XX P-PSDB; AAB56582.
XX
XX Prostate cancer associated gene sequences, referred to as prostate
XX cancer antigens, useful for treatment, prevention, and diagnosis of

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 25, 2003, 08:10:26 ; Search time 3972 Seconds

(without alignments)
2852.962 Million cell updates/secTitle: US-09-896-522-2
Perfect score: 1450
Sequence: 1 MASAGEDESPAPFADRP.....HPGMLTSGKSHLESSRPH 277Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 288871 segs, 2045481336 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame+g2n.model -DEV=xlh
-Q/cg2_1/USPTO/US09896522/runat_21112003_184353_5003/app_query.faeta_1.455
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blisum62 -TRANS=human40.cdd -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pro -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09896522.QCGN_1.1.3508.0/runat_21112003_184353_5003 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:*
1: gb_da:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pi:*
9: gb_pr:*
10: gb_ro:*
11: gb_stb:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_da:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pac:*
24: em_ph:*
25: em_pi:*
26: em_ro:*
27: em_stb:*
28: em_un:*

29: em_vl:*
30: em_hcg_hum:*
31: em_hcg_inv:*
32: em_hcg_other:*
33: em_hcg_mus:*
34: em_hcg_pin:*
35: em_hcg_rod:*
36: em_hcg_mam:*
37: em_hcg_vrt:*
38: em_sy:*
39: em_hcgo_hum:*
40: em_hcgo_mus:*
41: em_hcgo_other:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1450	100.0	834	6	AX449219	AX449219 Sequence
2	1450	100.0	834	9	AF254133	AF254133 Homo sapi
3	1450	100.0	1022	9	AF237290	AF237290 Homo sapi
4	1450	100.0	1624	6	AX449217	AX449217 Sequence
5	1450	100.0	2160	6	BD157613	BD157613 Primer to
6	1450	100.0	2160	6	AK022317	AK022317 Homo sapi
7	1387.5	95.7	2228	9	AK057848	AK057848 Homo sapi
8	1372.5	94.7	1395	9	AF125106	AF125106 Homo sapi
9	1325	91.4	1959	10	BC025146	BC025146 Mus muscu
10	1266	87.3	1810	10	MUSURKI	L31783 Mus musculu
11	1239	85.4	2072	9	BC015547	BC015547 Homo sapi
12	1122	83.6	254933	2	AC078885	AC078885 Mus muscu
13	1159.5	80.0	952	10	MUSURKIA	L31784 Mus musculu
14	1076	74.2	192336	2	AC118474	AC118474 Mus muscu
15	1076	74.2	211075	10	AC076974	AC076974 Mus muscu
16	1074.5	74.1	753	6	BD146824	BD146824 Primer to
17	941.5	64.9	1209	9	BC002906	BC002906 Homo sapi
18	941.5	64.9	1322	6	AX135546	AX135546 Sequence
19	934	64.4	1121	9	AF236637	AF236637 Homo sapi
20	934	64.4	1312	10	BC023789	BC023789 Mus muscu
21	934	64.4	1354	10	AF236636	AF236636 Mus muscu
22	932.5	64.3	744	9	AB062451	AB062451 Homo sapi
23	924	63.7	734	6	AX540411	AX540411 Sequence
24	909.5	62.7	1784	5	BC045968	BC045968 Danio rer
25	733	50.6	1467	3	AY119583	AY119583 Drosophil
26	656.5	45.3	147492	9	AL358781	AL358781 Human DNA
27	656.5	45.3	241882	10	AL808027	AL808027 Mouse DNA
28	656.5	45.3	254677	2	AC068494	AC068494 Mus muscu
29	631.5	43.6	244105	2	AC098897	AC098897 Rattus no
30	622.5	42.9	156905	2	AC020006	AC020006 Drosophil
31	622.5	42.9	199016	3	AC008204	AC008204 Drosophil
32	622.5	42.9	206741	3	AE003747	AE003747 Drosophil
33	572.5	39.5	136240	3	AC117070	AC117070 Dictyoste
34	570.5	39.3	1492	8	BT002336	BT002336 Arabidops
35	570.5	39.3	1844	8	AY089970	AY089970 Arabidops
36	545.5	37.6	1847	6	AX463241	AX463241 Sequence
37	545.5	37.6	1847	9	AK000524	AK000524 Homo sapi
38	544	37.5	1401	6	AX507248	AX507248 Sequence
39	544	37.5	1432	8	AY122946	AY122946 Arabidops
40	544	37.5	1785	8	AY072218	AY072218 Arabidops
41	542	37.4	3246	3	AY129436	AY129436 Drosophil
42	540.5	37.3	1813	6	AX405842	AX405842 Sequence
43	540.5	37.3	1875	9	BC033078	BC033078 Homo sapi
44	539.5	37.2	1823	10	BC016535	BC016535 Mus muscu
45	529	36.5	1831	8	AY086133	AY086133 Arabidops

RESULT 1

ALIGNMENTS

AX449219
 LOCUS AX449219 834 bp DNA linear PAT 03-JUL-2002
 DEFINITION Sequence 3 from Patent WO0202761.
 ACCSSION AX449219
 VERSION AX449219.1 GI:21697996
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 1 Gluckmann, M.A.
 57658, a human uridine kinase and uses thereof
 JOURNAL Patent: WO 0202761-A 3 10-JAN-2002;
 Millennium Pharmaceuticals, Inc. (US)
 Location/Qualifiers
 source 1..834
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"

BASE COUNT 194 a 218 c 265 g 157 t

ORIGIN

Alignment Scores:
 Pred. No.: 4.11e-129 Length: 834
 Score: 1450.00 Matches: 277
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 Gaps: 0

US-09-896-522-2 (1-277) x AX449219 (1-834)

QY 1 MetAlaSerAlaGlyGlyGluAspCysGluSerProAlaProGluAlaAspArgProHis 20
 Db 1 ATGGCTTCGGCGGAGGCGAAGACTCGAGAGCCCGCGCGGAGCCGACCGTCCGCAC 60

QY 21 GlnArgProPheLeuIleGlyValSerGlyGlyThrAlaSerGlyLysSerThrValCys 40
 Db 61 CAGCGCCCTTCCTGATAGGGGTAGCGCGGCACTCGCAGCGGGAAGTCGACCGTGTGT 120

QY 41 GlnLysIleMetGluLeuGlyGlnAsnGluValGluGlnArgGlnArgLysValVal 60
 Db 121 GAGAGATCATGAGTGTCTCGGACAGACAGAGTGAACGCGCAGCGGAGGTGTCTC 180

QY 61 IleLeuSerGlnAspArgPheTyrLysValLeuThrAlaGluGlnLysValAlaLys 80
 Db 181 ATCTGAGCCAGCAGAGTTCCTCAAGGTCTCTGACGACAGCAGAGCCAGGCTTG 240

QY 81 LysGlyGlnTyrAsnPheAspHisProAspAlaPheAspAsnLeuMetHisArgThr 100
 Db 241 AAAGAGACAGTCAATTTTACCATCCAGATCCCTTTGATTAATGTTGATGACAGACT 300

QY 101 LeuLysAsnIleValGluGlyLysThrValGluValProThrTyrAspPheValThrHis 120
 Db 301 CTGAGAAACATCTCGTGAAGGCGGCAAAACGCTGAGAGTCCGACCTTATTTTGTACACAC 360

QY 121 SerArgLeuProGluThrThrValValTyrProAlaAspValValLeuPheGluGlyIle 140
 Db 361 TCAAGGTACACAGACACACGCTGTCTACCTCGGACGCGTGTCTGTTGAGGCGCATC 420

QY 141 LeuValPheTyrSerGlnGluIleArgAspMetPheHisLeuArgLeuPheValAspThr 160
 Db 421 TTGGTGTCTTACACCCAGAGATCCGGACATGTTTCCACCTGCGCTCTTGTGTGACACC 480

QY 161 AspSerAspValArgLeuSerArgValLeuArgAspValArgArgGlyArgAspLeu 180
 Db 481 GACTCCAGCTCAGGCTGTCTCGAAGAGTTCTCCGGACGTCGCCCGAGGAGGAGACTTG 540

QY 181 GlnGlnIleLeuThrGlnTyrThrThrPheValLysProAlaPheGluGluPheCysLeu 200
 Db 541 GAGGAGATTCTGACGACAGTACACACACTTGTGTAAGCGGCGCTTGAAGAGTTCTGCTG 600

QY 201 ProThrLysLysThrValAlaAspValIleIleProArgGlyValAspAsnMetValAlaIle 220
 Db 601 CCACCAAGAGATATCCAGATGATCATCCACGAGAGTGAACATATGTTGCCATC 660

QY 221 AsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysLysTrpHisArg 240
 Db 661 AACCTAGCTGTGACGACATCCAGACATTTCTGAATGGTGAATCGCAATGCCACCGA 720

QY 241 GlyGlySerAsnGlyArgSerTyrLysArgThrPheSerGluProGluAspHisProGly 260
 Db 721 GGAGGTCACAAATGGCGGAGCTCAAGCGGACCTTTCTGAGCCAGGAGCACCTGTGG 780

QY 261 MetLeuThrSerGlyLysArgSerHisLeuGluLysSerSerArgProHis 277
 Db 781 ATGCTAGCTCTTGCAACGGTCACTTTGAGATCCAGACAGACCCCAAC 831

RESULT 2
 AF254133
 LOCUS AF254133 834 bp mRNA linear PRI 02-MAY-2001
 DEFINITION Homo sapiens uridine kinase mRNA, complete cds.
 ACCSSION AF254133
 VERSION AF254133.1 GI:13924749
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 1 (bases 1 to 834)
 Ho, Y.S. and Johnson, R.K.
 Human uridine kinase from prostate cancer cell line (LNCap)
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 834)
 Ho, Y.S. and Johnson, R.K.
 Direct Submision
 Submitted (10-APR-2000) Oncology Research, SmithKline Beecham, 709
 Swedeland Road, King of Prussia, PA 19406, USA
 Location/Qualifiers
 source 1..834
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 1..834
 /note="protein kinase"
 /codon_start=1
 /product="uridine kinase"
 /protein_id="AAK49122.1"
 /db_xref="GI:13924750"
 /translation="MASAGDEDCESPAPPADRPHQRPFLIGVSGTASGKSTVCEKIM
 ELIGQNEVROBORKVYILSDRFYKVLTEOKAKALKQGNPDPAFPNDLMHRTLK
 NIVEGTVVADPTVTHSRLEPTTVYVADVLVPEGLILVFSOEIRDMFHLRLPVD
 DSDVRLSRVLDVDRGRDLEQLITQYTTFYVAPAEPEFLPKTKYADVITIRPGVDNV
 AINLIVQHIQDILINDICMKHNGSGNSGRYSKTFSEPPGHPMLTSGKSHLESSSRP
 H"

BASE COUNT 194 a 219 c 265 g 156 t

ORIGIN

Alignment Scores:
 Pred. No.: 4.11e-129 Length: 834
 Score: 1450.00 Matches: 277
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 Gaps: 0

US-09-896-522-2 (1-277) x AF254133 (1-834)

QY 1 MetAlaSerAlaGlyGlyGluAspCysGluSerProAlaProGluAlaAspArgProHis 20
 Db 1 ATGGCTTCGGCGGAGGCGAAGACTCGAGAGCCCGCGCGGAGCCGACCGTCCGCAC 60

QY 21 GlnArgProPheLeuIleGlyValSerGlyGlyThrAlaSerGlyLysSerThrValCys 40
 Db 61 CAGCGCCCTTCCTGATAGGGGTAGCGCGGCACTCGCAGCGGGAAGTCGACCGTGTGT 120

Qy	41	GLIULYSILEMETGLULEULEUGLYGLINANGULVALIGLUGLINARGGLINARGYSVALVAL	60
Db	121	GAGAAAGATCATGAGAGTTCTCTGGGACAGAACGAGGTGGAAACACGGCAGCGGAAAGGTGGTC	180
Qy	61	ILEUSERGINAASPARGPHEITYRYSVALLEUTHRALAGLUGLILYSALALALEU	80
Db	181	ATCTGCAGCCAGAGACAGGTTCTACAGGTCCTGACGCGACAGACAGAGGCCCAAGGCTTTG	240
Qy	81	LYSGLYGLINTYRAPHENPHEASPHISPROAPALAPHEASPMASPMLEUWECHISARGTHR	100
Db	241	AAAGGACAGTCAATTTGATCCATCAATGCGCTTTGACAAATGATTTGATGACAGAGACT	300
Qy	101	LEULYSANILILEVALIGLUGLYLYSETHRVALGIVUALPROTHITYRASPHEVALTHRHS	120
Db	301	CTGAAGAACATCTGTGGAGGGCAAAACGGTGGAGGGCGACCTAAGATTGTTGTGACACAC	360
Qy	121	SERARGLEUPROGLUTHRTHRVALVALTYRPROALAASPVALVALLEUPHEGLUGLYLE	140
Db	361	TCAAGGTTACCGAGAGACACACGGTGGTCTACCTCGCGAGCGTGGTCTGTGTGAGGGCATC	420
Qy	141	LEUVALPHEITYRSERGINGLULILEARGAPMETPHEHISLEUARGLEUPHEVALASPTHR	160
Db	421	TTGGGTGTTCAAGCCAGACAGATCCGGGACATGTTCCACCTGCGCTCTTCGTGACACAC	480
Qy	161	ASPSEARSPVALARGLEUSERARGRYVALLEUAGAAPVALARGARGGLIYARGSPLEU	180
Db	481	GACTCCGACGTACAGGCTGTCTCGAAGAGTTCTCCGGACGTGCGCGCGAGGAGGAGACCTG	540
Qy	181	GLUGLILILEUTHRGLINTYRTHRTHRPHIVALLYSPROALAPHEGLUGLUPHECYSEU	200
Db	541	GAGCAGATTCTGACCGCAGTACACACACTTCGTGAGACCGGCTTGAGAGAGTTCTGGCTG	600
Qy	201	PROTHRILYSLYRYRALASPVALILEILEPROARGLYVALASPAMWEVALALALE	220
Db	601	CCGACAAAGAGTATGCCGATGTGATCAATCCACACAGAGATGGACAATAATGTTGCCATC	660
Qy	221	ASMEULILEVALGINHISILEGLNAPSIILEUAMANGIYASPIILECYLYSTPRHISARG	240
Db	661	AACTGATCTGACACACACATCCAGGACATTCGAAATGGAGCATGTGCAAATGGACCCGA	720
Qy	241	GLIYGLYSEARAAGLIYARGSERTYRILYSARGTHRPHESERGLUPROGLIYASPHISPROGLY	260
Db	721	GGAGGGTCCAAATGGCGGAGGTACAGAGGGAGCCTTTCTGAGCCAGGGGACCACTCTGG	780
Qy	261	METLEUTHRSERGILYLYARGSERHISLEUGLUSERSERSEARGIPROHIS	277
Db	781	ATGCTGACCTCTGGCAAAACGGTCAATTTGGAGTCCACAGACAGACCCAC	831
RESULT 3	AF237290	1022 bp	mRNA linear PRI 02-AUG-2001
LOCUS	AF237290		
DEFINITION	Homo sapiens uridine-cytidine kinase 1 (UCK1) mRNA, complete cds.		
ACCESSION	AF237290		
KEYWORDS	AF237290.1 GI:13506764		
SOURCE			
ORGANISM	Homo sapiens (human)		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	Van Rompay,A.R., Norda,A., Linden,K., Johansson,M. and Karlsson,A.		
JOURNAL	Phosphorylation of uridine and cytidine nucleoside analogs by two		
MEDLINE	human uridine-cytidine kinases		
PUBMED	Mol. Pharmacol. 59 (5), 1181-1186 (2001)		
REFERENCE	11306702		
AUTHORS	(bases 1 to 1022)		
TITLE	Van Rompay,A.R., Linden,K., Norda,A., Zhu,C., Zheng,X.,		
JOURNAL	Johansson,M. and Karlsson,A.		
REFERENCE	Human uridine-cytidine kinase 1 and 2 : rate limiting enzymes		
	Unpublished		
	(bases 1 to 1022)		

FEATURES			
source	location/Qualifiers		
gene	1..1022	/organism="Homo sapiens"	
		/mol_type="mRNA"	
		/db_xref="taxon:9606"	
	1..1022		
	5..838	/gene="VUCK1"	
		/gene="VUCK1"	
		/function="phosphorylates uridine and cytidine to uridine monophosphate and cytidine monophosphate"	
		/codon_start=1	
		/product="uridine-cytidine kinase 1"	
		/protein_id="AAK8324.1"	
		/db_xref="GI:13506765"	
		/translation="MASAGDEDCSPAPEDRPHORBPFLIGVSGTASGKSTVECKIMELCONVEORKAVLISODRFYKATLAEGAKALGOYVDPHDPADFMRLTLKNSDKTAEVPTVDFVTHSRLEPTQYTTTFPKAPFEECLPTKXADVILIRGVDMVDSVDRISRVRVDRRRDRLEQILTQYTFPKAPFEECLPTKXADVILIRGVDMVATNIVVHIDIDLNGDICMKHGGSSNGSRSTYKRTSEPDHPGLTSGKSHLBSRR	
gene			
CDS			
BASE COUNT	227 a	282 c	324 g 189 t
ORIGIN			
Alignment Scores:			
Pred. No.:	5.33e-129	Length:	1022
Score:	1450.00	Matches:	277
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0
US-09-896-522-2 (1-277) x AF237290 (1-1022)			
QY	1	MetlaserlAaglygyluaspcygluserProalAprogluAlaasparProhis	20
Db	5	ATGGCTTGCGGGGAGGGGAAAGACTGCGAGAGCCCGCGGAGCGCGACGTCGCGAC	64
QY	21	GlnatgProPheleuilegilyValsergilyThralasergilylaserThralCys	40
Db	65	CAGCGGCCCTTCCTGATAGGGGTAGCGCGCGCACTCGCAGCGGGAAGTCGCGTGT	124
QY	41	GluylsileneetGluLeuengilyGlnasnilyValGlnargGlnarglyValVal	60
Db	125	GAGAAAGATCAAGAGATTGCTGGGACAGAACGAGGTGAAACGCGGACGGAAGTGATC	184
QY	61	IleleuserGlnaaspargPheThylylValleuThralagluGlnilylAlaAlaAla	80
Db	185	ATCTGAGCCCGAGGACAGGTTCTTAAAGGTCCTGACGCGACGAGAAAGGCCAAGGCTTG	244
QY	81	LysgylGlnlyrAsnPhaasphleProasphAlaPheasphaenaleuMetHisArgThr	100
Db	245	AAAGGACAGTCAATTGTCACCAATCCAGATGCTTTGATATGATTTGATGACACGACT	304
QY	101	LeuylsAsnileValGluGlylylThThValGluValProThryrAspPheValThrhIs	120
Db	305	CTGAAGAAACATCGTGAAGGCGAANAACGTGAGGTGCGACCTATGATTTGTGACACAC	364
QY	121	SertrgleuProgluThrThThValVallyrProAlaasphValleuPheGluGlylle	140
Db	365	TCAAAGTTTACAGAGACCAACGCTGTCTACCTTCGACGCTGTCTTGTGAAGGCAATC	424
QY	141	LeuValPheThyrsargGlnilylleArgAspPhePheHisLeuArgleuPheValAspThr	160
Db	425	TTGGTGTTCACACCCAGAGATCCGGGACATGTTCCACTTCGCGCTTCGTGAGACAC	484
QY	161	AspserAspValArgleuserArgArgValleuArgasphValArgargGlyArgAspLeu	180

```
DB 485 GACTCCGACGTCAGGCTGTCTCGAAGAGTTCTCCGGGACGTCGCCGAGGAGGAGCTG 544
QY 181 GlnGlnIleLeuThrGlnIleThrThrPheValIleProAlaPheGlnGluPheCysLeu 200
DB 545 GAGCAGATTCTGACGCGATACACACACTTCCTGGAAGCCGCGCTTCGAGGAGTTCTGCTG 604
QY 201 ProThrIleValIleValIleValIleValIleValIleValIleValIleValIleValIle 220
DB 605 CCGCAAGAAGATATGCGATGTATCTATCCCGGAGAGTGAACAATATGTTGCAATC 664
QY 221 AsnLeuIleValIleGlnIleIleGlnAspIleLeuAsnGlnIleCysIleValIleValIle 240
DB 665 AACCTGATCGTGACGACATCCAGGACATCTGATGATGTCATCTGCAATATGCGACCGA 724
QY 241 GlyGlySerAsnGlyIleGlySerIleValIleGlyThrPheSerGluProGlyAspHisProGly 260
DB 725 GAGAGGCTTCATATGCGGAGCTACAGAGGAGCACTTTCTGAGCCAGGAGGACCACTTGGG 784
QY 261 MetLeuThrSerGlyIleValIleGlySerHisIleuGluSerSerSerArgProHis 277
DB 785 ATGCTGACCTCTGGCAACGGTCAATTTGGAGTTCAGACAGACACCCAC 835

RESULT 4
AX449217 1624 bp DNA linear PAT 03-JUL-2002
LOCUS AX449217
DEFINITION Sequence 1 from Patent WO0202761.
ACCESSION AX449217
VERSION AX449217.1 GI:21697994
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
AUTHORS Glucksmann, M.A.
TITLE 57658, a human uridine kinase and uses thereof
JOURNAL Patent: WO 0202761-A 1 10-JAN-2002;
Millennium Pharmaceuticals, Inc. (US)
FEATURES
SOURCE Location/Qualifiers
1. 1624
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
94..927
/note="unamed protein product"
/codon_start=1
/protein_id="CAD37771.1"
/db_xref="GI:21697995"
/translation="MASAGSDCESPAEADRPHPRLIGVSGTASGKSTCEKIM
ELIGONEVEORORRVILISODRFYKVLTAEGKAKALGOVNFDPDAFNDIMHRLTK
NIVGKTEVEPTYPVTHLSRLPETTVYPPADVLPFGILVPSGEIDMFLHLPVDT
DSVRLSRVLRKDRGRLDQILIQYITTFKAPFEELPKYIAVAILPRGVDNV
AIIIVGHIDILINGDICMKHRRGSSNRSYKRPSEPDHGMLTSGKSHLESSRP
H"
CDS
BASE COUNT 354 a 427 c 506 g 337 t
ORIGIN
Alignment Scores:
Pred. No.: 9.65e-129 Length: 1624
Score: 1450.00 Matches: 277
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0
US-09-896-522-2 (1-277) x AX449217 (1-1624)
QY 1 MetAlaSerAlaGlyGlyGluAspCysGluSerProAlaProGlnAlaAspArgProHis 20
DB 94 ATGGCTTCGGCGGAGGCGAAGACTGCGAGAGCCCGCGCGAGCGACCGATCCGAC 153
QY 21 GlnArgProPheLeuIleGlyValIleSerGlyIleThrAlaSerGlyIleSerThrValCys 40
DB 154 CAGCGGCCCTTCCTGATAGGGGTGAGCGGGGCGACTGCCAGCGGAGATCGACCGTGTGT 213
QY 41 GlnIleIleMetGlnIleLeuGlnIleGlnAsnGlnValIleGlnIleArgGlnIleValIle 60
DB 214 GAGAAATATCATGATGATGTTGCTGGGACGAAAGAGTGAACAGCGGAGCGGAGGAGTGC 273
QY 61 IleLeuSerGlnAspArgPheIleValIleValIleValIleValIleValIleValIleVal 80
DB 274 ATCTGAGCCAGGACAGGTTCTCAAGTCTTACCGGACGAGACGAAAGGCCAAGGCTTGG 333
QY 81 LysGlyGlnIleValIleValIleValIleValIleValIleValIleValIleValIleVal 100
DB 334 AAGGACAGTACATTTTGACCATCCAGATGCTTGTATATGATTTGATGACAGAGCAT 393
QY 101 LeuIleAsnIleValIleGlnIleGlyIleValIleValIleValIleValIleValIleVal 120
DB 394 CTGAAGAACATCTGTGAGAGGCAAAAGGAGAGTGCACCTATATATTTTGACACAC 453
QY 121 SerArgLeuProGluThrThrValIleValIleValIleValIleValIleValIleValIle 140
DB 454 TCAAGTTACCAAGACACAGGTGTCTTACCTCGGAGCTGTCTTGTGAGGCGATC 513
QY 141 IleValPheIleSerGlnGlnIleValIleValIleValIleValIleValIleValIleVal 160
DB 514 TTGGTGTCTTACAGCGCAGGAGATCCGAGCATGTTCACCTGGCGCTTCTGAGGCGATC 573
QY 161 AspSerAspValIleGlySerIleValIleValIleValIleValIleValIleValIleVal 180
DB 574 GACTCCGACGTCAGGCTGTCTCGAAGAGTCTCCGAGAGCTGCGGAGGAGGAGGAGCTG 633
QY 181 GlnGlnIleLeuThrGlnIleThrThrPheValIleProAlaPheGlnGluIlePheCysLeu 200
DB 634 GAGCAGATTCTGACGACGATACACCACTTCGTGAACCGGCTTCAGAGAGTCTCTCCTG 693
QY 201 ProThrIleValIleValIleValIleValIleValIleValIleValIleValIleValIle 220
DB 694 CCGACAAAGAAGATGCCGATGTATCCACGAGAGTGAACAATATGTTGTCATC 753
QY 221 AsnLeuIleValIleGlnIleIleGlnAspIleLeuAsnGlyIleValIleValIleValIle 240
DB 754 AACCTGATCTGAGACACATCCAGGACATTCGAAATGTGCAATCGCAATGCGACCGA 813
QY 241 GlyIleSerAsnGlyIleGlySerIleValIleGlyThrPheSerGluProGlyAspHisProGly 260
DB 814 GAGGCTTCATATGCGGAGCTACAGCGGACCTTTCTGAGCCAGGAGGACCACTTGGG 873
QY 261 MetLeuThrSerGlyIleValIleGlySerHisIleuGluSerSerSerArgProHis 277
DB 874 ATCTGACCTCTGGCAACGGTCAATTTGGAGTTCAGACAGACACCCAC 924

RESULT 5
BD157613 2160 bp DNA linear PAT 17-JAN-2003
LOCUS BD157613
DEFINITION Primer for synthesizing full-length cDNA and use thereof.
ACCESSION BD157613
VERSION BD157613.1 GI:27863371
KEYWORDS JP 2002191363-A/12456.
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 2160)
AUTHORS Ota, T., Iwagaki, T., Nishikawa, T., Hayashi, K., Saito, K., Yamamoto, J.,
Ishii, S., Sugiyama, T., Makamatsu, A., Nagai, K. and Otsuki, T.
TITLE Primer for synthesizing full-length cDNA and use thereof
JOURNAL Patent: JP 2002191363-A 12456 09-JUL-2002;
HELIX RESEARCH INSTITUTE
OS Homo sapiens (human)
PN JP 2002191363-A/12456
PD 09-JUL-2002
PF 28-JUL-2000 JP 2000280990
PI TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU
SAITO,
```


Query Match: 100.00% Indels: 0
 DB: 9 Gaps: 0
 US-09-896-522-2 (1-277) x AK022317 (1-2160)

QY 1 MetAlaSerAlaGlyGlyGluAspCysGluSerProAlaProGluAlaAspArgProHis 20
 DB 95 ATGGCTTCGGCGGGAGCGGAAGACTGCGAGAGCCCGCGCGGAGCGGACCGTCCGAC 154
 QY 21 GlnArgProPheLeuIleGlyValSerGlyGlyThrAlaSerGlyLysSerThrValCys 40
 DB 155 CAGGGCCCTTCGATGAGGGGTGAGCGGCGGACCTGCGACGGGAACTCAACGTGTGT 214
 QY 41 GlnLysIleMetGluLeuLeuGlyGlnAsnGluValGluGlnArgGlnArgLysValVal 60
 DB 215 GAGAAAGTCAATGAGTCTCTGGACAGACAGAGGTGGAACAGCCGCGGAGGAGGTGTC 274
 QY 61 IleLeuSerGlnAspArgPheTyrlsValValLeuThrAlaGluGlnLysAlaLysAlaLeu 80
 DB 275 ATCTGAGCCAGGACAGAGGTCTCAACAGGTCTTGAACGGCAGAGCAGAAAGCCAGGCTTG 334
 QY 81 LysGlyGlnTyrraenPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThr 100
 DB 335 AAAGGACAGTCAATTTTGACCATTCAGATGCTTTGATTAATGATTTGATTCAGAGACT 394
 QY 101 LeuLysAsnIleValGluGlyLysThrValGluValProThrTyrraAspPheValThrHis 120
 DB 395 CTGAAGAACATCTCGTGAAGGCGAAACGTTGAGGTGCGGACCTTATGATTTGTGACACAC 454
 QY 121 SerArgLeuProGluThrThrValValTyrrProAlaAspValValLeuPheGluGlyIle 140
 DB 455 TCAAGGTACCAAGACACACCGGTGCTACCTCGGCGGAGTCTGTGTTGAGGAGCATC 514
 QY 141 LeuValPheTyrrSerGlnGluIleArgAspMetPheHisLeuArgLeuPheValAspThr 160
 DB 515 TTGGTGTCTACAGCAGCAAGATCCGGGACATGTTCCACCTGCGCTTTCTGTGAGACACC 574
 QY 161 AspPheAspValArgLeuSerArgArgValLeuArgAspValArgArgGlyArgAspLeu 180
 DB 575 GACGCCGACGTCAGGCTCTCTCGAAGAGTCTCCGGGACGTGCCCGGAGGAGGAGCCTG 634
 QY 181 GluGlnIleLeuThrGlnTyrrThrThrPheValLysProAlaPheGluGluPheCysLeu 200
 DB 635 GAGCAGATTCGAGCGCAGTACACACCTTCGTGAAGCCGCGCTTCGAGAGTTCGTGCTG 694
 QY 201 ProThrLysLysTyrrAlaAspValIleIleProArgGlyValAlaAspAsnMetValIle 220
 DB 695 CCGACAAAGATGATGCCGATGTGATCATCCACGAGAGTGAACAATATGTTGCCATC 754
 QY 221 AsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysLysTyrHisArg 240
 DB 755 AACCTGATCGTGCAGCAGCATCCAGGACATTCGAAATGTGTACATCTGAAATGTGCACCA 814
 QY 241 GlyLysSerAsnGlyArgSerTyrrLysArgThrPheSerGluProGlyAspHisProGly 260
 DB 815 GAGAGGCTCCATGTGGCGGAGCTACAGCGGACCTTTCTGAGCCAGGGGACACCCCTGG 874
 QY 261 MetLeuThrSerGlyLysArgSerHisLeuGluLysSerSerArgProHis 277
 DB 875 ATGCTGACCTCTGGCAACGGGTCAACATTTGAGTCCAGACAGACACCCAC 925

RESULT 7
 AK057848 2228 bp mRNA linear PRI 31-OCT-2001
 LOCUS AK057848
 DEFINITION Homo sapiens cDNA FLJ25119 fis, clone CBR05878, highly similar to
 URIDINE KINASE (EC 2.7.1.48).
 ACCESSION AK057848
 VERSION AK057848.1 GI:16553809
 KEYWORDS oligo cloning; fis (full insert sequence).
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 AUTHORS
 1 Tashiro,H., Yamazaki,M., Watanabe,K., Kumagai,A., Itakura,S.,
 Fukuzumi,Y., Fujimori,Y., Komiyama,M., Suzuki,Y., Hata,H.,
 Nakagawa,K., Mizuno,S., Morinaga,M., Kawamura,M., Sugiyama,T.,
 Irie,R., Otsuki,T., Sato,H., Nishikawa,T., Sugiyama,A.,
 Kawakami,B., Nagai,K., Isegai,T., and Sugano,S.
 NEDO human cDNA sequencing project
 JOURNAL
 TITLE
 NEDO human cDNA sequencing project
 REFERENCE
 2 (bases 1 to 2228)
 AUTHORS
 Sugano,S. and Suzuki,Y.
 TITLE
 Direct Submission
 JOURNAL
 Submitted (24-OCT-2001) Sumio Sugano, Institute of Medical Science,
 University of Tokyo, Laboratory of Genome Structure, Human Genome
 Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan
 (E-mail:cdna@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
 Fax:81-3-5449-5416)
 COMMENT
 NEDO human cDNA sequencing project supported by Ministry of
 Economy, Trade and Industry of Japan; cDNA full insert sequencing:
 Research Association for Biotechnology (RAB); cDNA library
 construction and 5'-end one pass sequencing: Institute of Medical
 Science, University of Tokyo, Laboratory of Genome Structure, Human
 Genome Center; 3'-end one pass sequencing: RAB; clone selection for
 full insert sequencing: RAB and Helix Research Institute.
 FEATURES
 Location/Qualifiers
 1..2228
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CBR05878"
 /issue_type="brain"
 /clone_lib="CBR"
 /note="Cloning vector: pME18SFL3"
 BASE COUNT 478 a 598 c 696 g 456 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 1.33e-122 Length: 2228
 Score: 1387.50 Matches: 276
 Percent Similarity: 88.75% Conservative: 0
 Best Local Similarity: 88.75% Mismatches: 1
 Query Match: 95.69% Indels: 35
 DB: 9 Gaps: 2
 US-09-896-522-2 (1-277) x AK057848 (1-2228)
 QY 1 MetAlaSerAlaGlyGlyGluAspCysGluSerProAlaProGluAlaAspArgProHis 20
 DB 64 ATGGCTTCGGCGGGAGCGGAAGACTGCGAGAGCCCGCGCGGAGCGGACCGTCCGAC 123
 QY 21 GlnArgProPheLeuIleGlyValSerGlyGlyThrAlaSerGlyLys----- 36
 DB 124 CAGCGGCCCTCTCTGATAGGGGTGAGCGCGGCGGACCTGCCAGCGGAGACCGCGGTTTC 183
 QY 37 -----SerThrValCysGluLysIleMetGluLeu 46
 DB 184 CAGGCGGGAATCCCGCTTCTGTCTTCCAGTCCAGCCGTGTGTGAAGAATCAAGAGTTG 243
 QY 47 LeuGlyGlnAsnGluValGluGlnArgGlnArgLysValValIleLeuSerGlnAspArg 66
 DB 244 CTGGAGACAGAACAGAGTGGAACAGCGGAGCGGAGAGGTGTCTCTGAGCCAGGACAGG 303
 QY 67 PheTyrrLysValLeuThrAlaGluGlnLysAlaLysAlaLeuLysGlyGlnTyrraAspPhe 86
 DB 304 TTCTACAGGTCTCTGACGCGAGGACGAGGAGCCAAAGCCCTTGAAAGACAGTCAATTTT 363
 QY 87 AspHisProAspAlaPheAspAsnAspLeuMetHisArgThrLeuLysAsnIleValGlu 106
 DB 364 GACCATTCAGATGCTTTGATTAATGATTTGATGATCAACAGGACTCTGAAGAACATCGTGAG 423
 QY 107 GlyLysThrValGluValProThrTyrraAspPheValThrHisSerArgLeuProGluThr 126
 DB 424 GGCAAAACGGTGAAGGTGCGGACCTATGATTTTGTGACACATCAAGGTTTACAGAGACC 483

Qy 127 ThrValValTyrProAlaAspValValIleuPheGluGluIleValAlaPheTyrSerGln 146
 Db 484 ACGGTGCTTACCTCGCGAGAGTGTCTGTTTAAAGGACATCTTGTTCTTACAGCCAG 543
 Qy 147 GluIleArgAspMetPheHISleuArgLeuPheValAspThrAspSerAspValArgLeu 166
 Db 544 GAGATCCGGGACATGTTCCACCTGCGCTCTTCGTGGACACGACTCGAGCTGAGGCTG 603
 Qy 167 SerArgArgValIleuArgAspValArgArgGlyValArgAspLeuGluGlnIleLeuThrGln 186
 Db 604 TCTCGAAGAGTCTCCCGGAGCTGCGCGAGGGAGGACCTGGAGCAGATTCTGACCGCAG 663
 Qy 187 TyrThrThrPheValIleuPheGluGluIlePheGluGluPheCysLeuPheProThr 202
 Db 664 TACACCACTTCCTGTAAGCCGCGCTTGGAGAGTTCCTGCTGCCCA-GAGCATCTGACA 722
 Qy 203 -----LysIleTyrAla 206
 Db 723 GGGAAATGAGAGTCAGCATTTGAGCCATGATGATTGATGAGGAGACAAAGAGTATGCC 782
 Qy 207 AspValIleIleProArgGlyValAspAsnMetValAlaIleAsnLeuIleValGlnHis 226
 Db 783 GATGTGATCATCCACAGAGAGTGGACATATGTTGCCATCAACTGATCTGACAGCAC 842
 Qy 227 IleGlnAspIleLeuAsnGlyAspIleCysIleTyrPheIleArgGlyIleSerAsnGlyValArg 246
 Db 843 ATCCAGAGACATTCGAATGTGATCTGCAATATGCAACGAGAGGAGTCCATATGGCGG 902
 Qy 247 SerTyrIleArgThrPheSerGluProGlyAspHisProGlyMetLeuThrSerGlyLys 266
 Db 903 AGCTACAGACGAGACTTTTCTGAGCCAGGGGACCACTCGGATGTCGACTCTGCGAAA 962
 Qy 267 ArgSerHisLeuGluSerSerSerArgProHis 277
 Db 963 CGGTCACTTTGGAGTCCAGACAGACCCAC 995

RESULT 8
 AF125106 1395 bp mRNA linear PRI 07-FEB-2002
 LOCUS Homo sapiens uridine kinase mRNA, complete cds.
 DEFINITION AF125106
 ACCESSION AF125106 GI:18568108
 VERSION AF125106.1 GI:18568108
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 1395)
 Xing,Y.R., Yu,L. and Zhao,S.Y.
 TITLE Cloning of a new human cDNA similar to Mus musculus uridine kinase mRNA
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 1395)
 Authors Ding,J.B., Yu,L. and Zhao,S.Y.
 TITLE Direct Submission
 JOURNAL Submitted (02-FEB-1999) Lab of Human Gene Research, Institute of Genetics, Pudan University, No. 220 Handan Rd., Shanghai 200433, People's Republic of China
 FEATURES
 source
 1. 1395
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 8. 838
 /codon_start=1
 /product="uridine kinase"
 /protein_id="AA175943.1"
 /db_xref="GI:18568109"
 /translation="MASAGSEGARARAGANRPHORPFLIGSGTASGKSTVCEKIME
 LIGQNEVEQHGKRVILSDRFYKVLAEAKALAKQVFPDPAFDMDLMRLLKN
 IVGKTVPEVPTDFVTHSRLEPTVVYPADVAFEGILVYSQEIRDMFRLRFVDT
 SDVLRVRVLRGRLDEQLTOYTFVKKPAFEERCLPTKYLAVIIPRGVNNVA
 INIIVHIDILINDICKMHRGSGNGRTYKRTSEPRDHDGMLTSGRSHLESSRPH

BASE COUNT 325 a 352 c 411 g 307 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 1,97e-121 Length: 1395
 Score: 1372.50 Matches: 270
 Percent Similarity: 98.19% Conservative: 2
 Best Local Similarity: 97.47% Mismatches: 5
 Query Match: 94.66% Indels: 3
 DB: Gaps: 0
 US-09-896-522-2 (1-277) x AF125106 (1-1395)
 Qy 1 MetAlaSerAlaGlyGlyAspCysGluSerProAlaProGluAlaAspArgProHis 20
 Db 8 ATGCTTCCGCGGAGAGCGAAGG-TGCGAGAG-CCGCGCGGAGAGG-AACTCTCCGAC 64
 Qy 21 GlnArgProPheLeuIleGlyValSerGlyIleThrAlaSerGlyLysSerThrValCys 40
 Db 65 CAGCGTCCCTTCTGATAGGGGTAGCGGGGACCTGCCAGCGGAGATCGACCTGTGT 124
 Qy 41 GluLysIleMetGluLeuLeuGlyGlnAsnGluValGluGlnArgGlnArgLysValAla 60
 Db 125 GAGAAGATCATGAGATTGCTGGGACAGAACAGGTGGAACAGCGGACGGTATGTGTC 184
 Qy 61 IleLeuSerGlnAspArgPheTyrLysValLeuThrAlaGluGlnLysAlaLysAlaLeu 80
 Db 185 ATCTGAGCCAGAGACAGTTCCTCAAGGTCTGACGCGAGACAGAACAGCGCTTG 244
 Qy 81 LysGlyGlnTyrTrpAsnPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThr 100
 Db 245 AARGAGACGTACATTTTGACATCCAGAGTCTTGATATGATTTGATGACAGAGACT 304
 Qy 101 LeuLysAsnIleValGluGlyLysThrValGluValProThrTyrAspPheValIleHis 120
 Db 305 CTGAAGAACATCGTGGAGGGCAAAACGTGGAGGTCCGACCTATGATTTTGTGACACAC 364
 Qy 121 SerArgLeuProGluThrThrValValIleTyrProAlaAspValAlaLeuPheGluGlyIle 140
 Db 365 TCAGGCTTACAGAGACCAAGGTGTCTACCTGCGAGAGCTGGTCTGTTGAGGCGATC 424
 Qy 141 LeuValPheTyrSerGlnGluIleArgAspMetPheHISleuArgLeuPheValAspThr 160
 Db 425 TTGCTTCTTACAGCGAGAGATCCGGGACAGTTCACCTGCGCTTCTTGAGCACACC 484
 Qy 161 AspSerAspValArgLeuSerArgValIleuArgAspValArgArgGlyArgAspLeu 180
 Db 485 GATTCGAGCTACAGCTGTCTCGAAGAGTTCCTCGGAGAGTGGCGCGGAGGAGACTG 544
 Qy 181 GluGlnIleLeuThrGlnTyrThrThrPheValLysProAlaPheGluGluPheCysLeu 200
 Db 545 GAGCAATTCGACGAGTACACACACACTTCGTGAACCGGCTTCGAGAGATTTCTCCCTG 604
 Qy 201 ProThrLysLysThrLysAspValIleIleProArgGlyValAspAsnMetValAlaIle 220
 Db 605 CCAGCAAAAGAGATGCCATGTGATCATCCACGAGAGTGACAAATATGTTGCCATC 664
 Qy 221 AsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysLysTyrPheIleArg 240
 Db 665 AACCTGATGTGAGACATCCAGAGCATTTCTGAATGTGACATCTGCAAAATGGACCGCA 724
 Qy 241 GlyGlySerAsnGlyValArgSerTyrLysArgThrPheSerGluProGlyLysAspHisProGly 260
 Db 725 GAGAGTCCCAATAGGGGAGAGTACAAAGCGAGCTTTCTGAGCCAGAGGAGACCACTCGGG 784
 Qy 261 MetLeuThrSerGlyLysArgSerHisLeuGluLeuSerSerSerArgProHis 277
 Db 785 ATGTCACTCTGGCAACGCTACATTGTGAGTCCAGACAGACAGACCCAC 835

RESULT 9
 BC025146 1959 bp mRNA linear ROD 16-APR-2003
 LOCUS BC025146

DEFINITION Mus musculus uridine monophosphate kinase, mRNA (cDNA clone
MGC:36231 IMAGE:4913412), complete cds.
ACCESSION BC025146
VERSION BC025146.1 GI:19263563
KEYWORDS MGC.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1959)
AUTHORS Strauberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shennen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Heien,F.,
Diatchenko,L., Marsina,K., Farmer,A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Scheetz,T.E., Brownstein,M.J., Uedini,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullahy,S.D., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gnaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Heiton,E., Kettelman,W., Madan,A., Rodriguez,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalek,U., Smalins,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
2238257
12477932
2 (bases 1 to 1959)
AUTHORS Strauberg,R.
TITLE Direct Submission
JOURNAL Submitted (05-MAR-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK
COMMENT NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: amg@bcm.tmc.edu
Gunnarone, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Louisgeed, H.,
Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nananavi,
A.N., Gibbs, R.A.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAX Plate: 61 Row: P Column: 19
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis, Similarity but not identity to protein.
Location/Qualifiers
1..1959
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="MGC:36231 IMAGE:4913412"
/tissue_type="Salivary gland, 10 week old female mouse"
/clone_lib="NCI CGAP_SG2"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
1..1959
gene

CDs
/gene="Umpk"
/db_xref="LocusID:22245"
/db_xref="MGI:98904"
57..890
/codon_start=1
/product="Umpk protein"
/protein_id="AAH25146.1"
/db_xref="GI:19263564"
/db_xref="LocusID:22245"
/translation="MAGAGGGGSESAAPENDRQPPPLIGVSGGTASGKSYCEIM
ELIQNEVDRQRKVLISQDCFYKLTAEQAKAKGQNFDPHAPDNDLHKTLK
NIVEGTSVEPTVDFVTHSRLEPTTVVADVVFEGILVFYQELRDMFHLFVDT
DSVRLSRVLRVQGRDLQILQITVTAFFKPAFEFCLPTKYADVII PRGVDMV
AIVLIYQHIQDILNGDLCKRHRRGPNRHNKRTFPPBDHPGLATGKRSLESSRP
H"
BASE COUNT 466 a 511 c 561 g 421 t
ORIGIN
Alignment Scores:
Pred. No.: 1,03e-116 Length: 1959
Score: 1325.00 Matches: 254
Percent Similarity: 95.67% Conservative: 11
Best Local Similarity: 91.70% Mismatches: 12
Query Match: 91.38% Indels: 0
Gaps: 0
US-09-896-522-2 (1-277) x BC025146 (1-1959)
QY 1 MetAlaSerAlaGlyGlyGlyAspCysGlySerProAlaProGluAlaAspArgProHis 20
Db 57 ATGGCTTCGGCGGAGGCGGCGGCGGCGGAGAGCGCGCGCCGAGCGGCGGATCGTCCCGAG 116
QY 21 GluArgProPheLeuIleGlyValSerGlyGlyThrAlaSerGlyLysSerThrValCys 40
Db 117 CCGGGCCGTTCTCATCGCGGTGAGGGGGGAGCCGCTAGTGCAAGTCAAGTCAAGTGTG 176
QY 41 GluLysIleMetGluLeuLeuGlyGlnAsnGluValGluGlnArgGlnArgLysValVal 60
Db 177 GAGAGAGTACGTGAGGCGTGGGAGACAGAACGAGTGAACCGCGGCGGCGCAAGTTGGTC 226
QY 61 IleLeuSerGlnAspArgPheIleLysValLeuThrAlaGluGlnLysAlaLysAlaLeu 80
Db 237 ATCTTGAGCCAGACATGCTTCTCAAGGTTCTGACCGCGGAGAGCGGCGGCGGCTTTG 296
QY 81 LysGluGlnIleThrAsnPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThr 100
Db 297 AAGGACAGTACATTTTGACCCACCAAGATGCTTTGATATATGATCTGATGCACAGACC 356
QY 101 LeuLysAsnIleValGluGlyLysThrValGluValProThrTyraAspPheValThrHis 120
Db 357 CTGAAAACATGTTTGAAGCCAAACGTCGAGAGTCCCTACCTATGATTTTGTGACCCAC 416
QY 121 SerArgLeuProGluThrThrValValIleTyroAlaAspValValLeuPheGluGlyIle 140
Db 417 TCAAGGTTACCAAGACCACTGTGGTCTACCCAGCTGATGTGGTGTCTGCGAGGCGATC 476
QY 141 LeuValPheTyroSerGlnGluIleArgAspMetPheHisLeuArgLeuPheValAspThr 160
Db 477 TTGATTTCTACACCCAGAGATCCGGGACATTTTCACTGGCGCTTTTGTGGACACA 536
QY 477 TTTGATTTCTACACCCAGAGATCCGGGACATTTTCACTGGCGCTTTTGTGGACACA 536
QY 161 AspSerAspValArgLeuSerArgValLeuArgAspValArgArgGlyArgAspLeu 180
Db 537 GACTCTGATGTGAGGCTGTCCAGAGGTTCTCCGGGAGTGCACAGAGAGAGGACCTG 596
QY 181 GluGlnIleLeuThrGlnIleThrThrPheValLysProAlaPheGluGluPheCysLeu 200
Db 597 GACGAGTCTGACTCAGTACACCGGCTTTGTGAAACCAAGCTTTAGAGGAGTCTGCTG 656
QY 201 ProThrLysLysThrAlaAspValIleIleProArgGlyValAspAsnMetValAlaIle 220
Db 657 CCGACTAAGAGTACGCTGAGTATCTCCCTCGAGAGGTTGATATATGAGGCGCATC 716
QY 221 AsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysLysThrPheArg 240

```

Db      717 AACGATGTCGGAACATCCAGCATCTCAAGGGGACCTGTCAAGCGGACCGA 776
Qy      241 G1YGI5eranglYarSerTYrLYsArYThrPheSerLupProG1YaPheS1pG1y 260
Db      777 GGGGGGCCCAAGCGGCGCAACAGAGACCTTCCCGACGAGAGATCAACCTGGG 836
Qy      261 MetLeuThrSerG1YsArSerPheS1eUGlUSeSerSerArpProH1s 277
Db      837 GTGTGGCCACTGGCAAGCGCTCACCTGGAGTCTAGACAGACCCCAT 887

RESULT 10
MUSURKI      1810 bp  mRNA  linear  ROD 27-MAR-1997
LOCUS        Mus musculus uridine kinase mRNA, partial cds.
ACCESSION    U11783
VERSION      U11783.1 GI:471980
KEYWORDS     Mus musculus (house mouse)
SOURCE       Mus musculus
ORGANISM     Mus musculus (house mouse)
REFERENCE    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE        Traut, T.W.
JOURNAL      Eur. J. Biochem. 222 (1), 9-19 (1994)
MEDLINE      94259063
PUBMED       8200357
REFERENCE    2 (bases 1 to 1810)
AUTHORS      Ropp, P.A. and Traut, T.W.
TITLE        Cloning and expression of a cDNA encoding uridine kinase from mouse
JOURNAL      Arch. Biochem. Biophys. 336 (1), 105-112 (1996)
MEDLINE      97108719
PUBMED       8951040
FEATURES     source
             location/Qualifiers
             1..1810
             /organism="Mus musculus"
             /mol_type="mRNA"
             /db_xref="taxon:10090"
             /tissue_type="brain"
             <1..783
             /EC_number="2.7.1.48"
             /standard_name="uridine-cytidine kinase"
             /product="uridine kinase"
             /protein_id="A850568.1"
             /db_xref="GI:471981"
             /translation="RROPRLGVSGGTASGKSTYCEKIMELGONEVDRQRKLYI
             LSODCFYKVLIAQKAKLKGQYNFDPDAFNDLMKTKLNIVSGKIVEVPTDFVT
             HSRLEPTTVVYPADVALFEGILVFTQELDMHMLFLPVTDSVLSRVLNDVORG
             RLEQILITQYTAKPAPFEFCFPTKKYADAVIIPRGDMMVAIINLIVQIHODILNGDL
             CKNRHGGPMGRNKRTPFBGDPHGVATATKSHLSSSRPH"
             join(37..63,412..426,451..465)
             /bound_moiety="ATP at catalytic site"
             misc_binding
             polyA_signal
             BASE COUNT 411 a 482 c 512 g 405 t
             ORIGIN
             polyA_signal
             1790..1795
             482 c 512 g 405 t

Alignment Scores:
Pred. No.:      3 97e-111      Length:      1810
Score:          1266.00      Matches:      241
Percent Similarity: 96.92%      Conservative: 11
Best Local Similarity: 92.69%      Mismatches: 8
Query Match:    87.31%      Indels:      0
DB:             10      Gaps:      0

US-09-896-522-2 (1-277) x MUSURKI (1-1810)
Qy      18 ArgProHisG1YsArSerPheS1eUGlUSeSerSerArpProH1s 37
Db      1 CGTCCAGCGGCGGCGGCTTCTCATCGCGGTAGCGCGGCAACCGCTAGTGAAGTCA 60

```

```

Qy      38 ThrValCYeGluuYs11eMeTg1UeLeuG1YInaSnG1UVal1UgInaRg1nArG 57
Db      61 ACAGTGTGAGAGATCATAGAGCTGCGGACAAAGAGTGAAGCGCGGACGCG 120
Qy      58 YsValVal11eUeSerG1nAspArPheTYrLYsVal1eUeThr1aG1UInYsAla 77
Db      121 AGTTGTATCTCTGAGCCAGAGCTCTTCAACAAGTTCTGACGGCCAGAGAGGCC 180
Qy      78 YsAl1eUeYsG1UInYrAnPheAsPheS1pRoAsP1aPheAsPheS1eUeMet 97
Db      181 AAGGCTTGAAGGAGACATCAATTTGACCAACCCAGATGCTTTGATATGATCGATG 240
Qy      98 H1dArGThrLeuYsSn11eValG1Ug1YsThYsValG1UVal1ProH1rYrAsPhe 117
Db      241 CAGPAGACCTGAAACATGTGTAAGGCAAACTGAGAGCTCCCTCACTATGATTTT 300
Qy      118 ValThH1SeSerArGLeuPProG1UTrThrVal1aThYrPro1aAsPVal1eUePhe 137
Db      301 GTACCCACTCAAGTTACCAAGACCACTGTGTCTACCCAGCTGATGGTTCTGTTT 360
Qy      138 G1Ug1Y1eUeYsAlPheTYrSerG1nG1U11eArGAsPhePheH1SeUeArGLeuPhe 157
Db      361 GAGGCGATCTGGTATTCTACACCCAGAGATCCGGAGATGTTCCACCTGCGCTCTTT 420
Qy      158 ValAsPThrAsPSerAsPValArGLeuSerArGArGVal1eUeArGAsPValArGArG1Y 177
Db      421 GTGGACACAGACTCATATGTAGCTGTCTCGAAGAGTTCTCGGATGTGCAACAGGA 480
Qy      178 ArgAsPLeuG1Ug1n1eUeThrG1nYrThrThrPheValYsPro1aPheG1Ug1U 197
Db      481 AGGAGACTGAGAGATCTGACTCGTACGTAACCGCTTTGTGAAACAGCTTTGAGAG 540
Qy      198 PheCYeLeuPProThrYsLYsTYrAlaAsPVal11e1eProArG1YsValAsPAsnMet 217
Db      541 TTCTGCTGCCGACCTAAGAAATAGCTGACGTGATCATCTCGAAGAGATTGTAATATG 600
Qy      218 ValAl1eAsnLeu11eValG1nH1SeG1nAsP11eUeAsnG1YsP11eCYeYs 237
Db      601 GTGGCCATCAACCTGATGTGCAACATCCAGAGATCTCAACGGGACCTGTCAAG 660
Qy      238 TrpH1eArG1YG1YSeranG1YArSerTYrLYsArYThrPheSerLupProG1YaP 257
Db      661 CGGACCGAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 720
Qy      258 H1sProG1YMetLeuThrSerG1YsArSerPheS1eUGlUSeSerSerArpProH1s 277
Db      721 CACCTGGGGGTGTGGCCACTGCGCAAGCGCTCACACTGGAAGTCTAGACAGACCCCAT 780

RESULT 11
BC015547      2072 bp  mRNA  linear  PRI 04-OCT-2001
LOCUS        Homo sapiens, Similar to uridine-cytidine kinase 1, clone MGC:9668
DEFINITION   IMAGE:3845821, mRNA, complete cds.
ACCESSION    BC015547
VERSION      BC015547.1 GI:15930229
KEYWORDS     MGC.
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
REFERENCE    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS      Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE        1 (bases 1 to 2072)
JOURNAL      Strausberg, R.
DIRECT SUBMISSION
Submitted (01-OCT-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK       NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT      Contact: MGC help desk
              Email: cgaabs-r@mail.nih.gov
              Tissue Procurement: ATCC
              cDNA Library Preparation: Life Technologies, Inc.

```

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome
 Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: <http://www-shgc.stanford.edu>
 Contact: (Dickson, Mark) mcdpaxil.stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
 R. M.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
 Series: IRAC Plate: 20 Row: 1 Column: 5
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 13899252.

FEATURES

SOURCE

1..2072
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:9668 IMAGE:3845821"
 /tissue_type="Colon, adenocarcinoma"
 /clone_lib="MGC_65"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"
 87..692
 /codon_start=1
 /product="similar to uridine-cytidine kinase 1"
 /protein_id="AAH15547.1"
 /db_xref="gi:15930230"
 /translation="MASAGGEDCESPAEADPHQPPFLIGVSGTASGKSYCEKIM
 ELIQNEVEORQKRVILSDRFKVLTAEBKAKALQGVNFDHPDAFNDLMHTLK
 NIVEGKIVEPTVDFVTHSRLEPTVVYPAVAVFEGILVFEYSQEIRDMFHLRLFVDT
 DSDVRLSRDRKVCRCDHPTSRSGOYGGQHPRAHPHSEM"

CDS

BASE COUNT 455 a 559 c 637 g 421 t
 ORIGIN

Alignment Scores:
 Pred. No.: 1.78e-108 Length: 2072
 Score: 1239.00 Matches: 244
 Percent Similarity: 87.77% Conservative: 0
 Best Local Similarity: 87.77% Mismatches: 1
 Query Match: 85.45% Indels: 33
 DB: 9 Gaps: 1

US-09-896-522-2 (1-277) x BC015547 (1-2072)

QY 1 Metlaseeraiaaglygluaapcyeguserproalaaprogluaaasparprohis 20
 Db 87 ATGGCTTCGGGGAGGAGGAGCTGCGAGAGCCCGCGGAGGCGGACCTCCGAC 146
 QY 21 Glnatgpprophenleuilegylvaliserclgltthralaserglylvsertthvalcys 40
 Db 147 CAGCGGCCCTTCCTGATRGSGGTAGCGCGCGGACCTGCGAGGAGAGTCCAGCTGTGT 206
 QY 41 Glulysilemetgluleuileuglyglinaangluvaluglnatrglnatrglyvala 60
 Db 207 GAGAAAGATCATGAGTGTCTGCGACAGAAACGAGGTGAAACGCGGACGCGAAGTGTGTC 266
 QY 61 Ileusererginaapargphetyrlylvalleuthralagluinlysalalyalaleu 80
 Db 267 ATCTGAGCCAGGACAGGCTTCAAGGTCTGACGCGAGGAGGAGGAGGAGGAGGCTTG 326
 QY 81 Lysgllyglntryraanpheasphileproasphalapheasphaaspleamethisargthr 100
 Db 327 AAAAGACAGTACATTTTGACATCCAGATCCCTTTGATTAATGATTTTGATGACAGAGACT 386
 QY 101 Leuylasaniilevaliguglylyserthvalgluvalprothtryraasphvalthhis 120
 Db 387 CTGAAGAACATCGTGAAGGCGAAACGCTGAGGCGGACCTATGATTTTGACACAC 446
 QY 121 Sertrgleuprogiluththvalvaltyrproalaasphvalleuphegluglylyle 140
 Db 447 TCAAGGTTACAGAGACCAAGGTGTCTACCTTCGCGAGCGTGTCTGTTGAGGAGCATC 506

QY 141 LeuValPheTyrSerGlnGluIleArgAspMetPheHisLeuArgLeuPheValAspThr 160
 Db 507 TTGGTGTCTTACAGCCAGAGATCCGGACATGTTCCACTGCGCTTCCTGCGAGACC 566
 QY 161 AspSerAspValArgLeuSerArg-ArgValLeuArgAspValArgArgGlyArgAspLe 180
 Db 567 GACTCCGACGTACAGGCTGTCTCGAAGAG----- 594
 QY 180 uGluGlnIleLeuThrGlnTrpThrPheValysProalaPheGluGluPheCysLe 200
 Db 594 ----- 594
 QY 200 uProThrLysLysTyrAlaAspValIleleProArgGlyValAspAspMetValaI 220
 Db 595 -----ACAAAGATGATCCGATGTGATCATCCACAGAGATGACAAATGTGCTCCAT 650
 QY 220 eAsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysLysTrpHisar 240
 Db 651 CAACCTGATCGTCAGACATCCAGACATTTGAAATGGTGACATCTGCMAATGGCACCG 710
 QY 240 ggIyGlySerAnglyArgSerTyrlysaatgthrpheaserGluProGlyAspHisProG 260
 Db 711 AGAGAGGTCCAAAGGCGGAGCTACAGACGACCTTTTCTGAGCCAGGGGACCACTCG 770
 QY 260 yMetLeuThrSerGlyLysArgSerHisLeuGluSerSerSerArgProHis 277
 Db 771 GATGCTGACTCTGCGAAAGCTCACATTTGAGTGTCCAGCAGACACCCAC 822

RESULT 12
 AC078885/c 254993 bp DNA linear HTG 15-MAY-2002
 LOCUS M5 musculus chromosome 6 clone RP3-117723 strain C57BL/6/J,
 DEFINITION WORKING DRAFT SEQUENCE, 44 unordered pieces.
 AC078885
 AC078885.8 GI:18376842
 VERSION HTG; HTGS PHASE1; HTGS DRAFT.
 KEYWORDS M5 musculus (house mouse)
 SOURCE M5 musculus
 ORGANISM Mus musculus

REFERENCE
 AUTHORS Montgomery, K.T., Grills, G., Han, J., Lee, E., Long, J., Pomerantz, R.,
 Ioshikhes, I.P., Shim, C., Decker, J., Thomas, E., Perera, A.,
 Gordon, M., Goltz, J.S. and Kucherlapati, R.
 High Throughput Mouse Sequencing
 Unpublished
 2 (bases 1 to 254993)
 JOURNAL
 REFERENCE
 AUTHORS

TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 COMMENT

Center: Harvard Partners Genome Center
 Center Code: HPGC
 Web site: <http://www.hpcg.org/sequence/mouse.html>
 Contact: hpcg@mednet.mgh.harvard.edu
 -----Summary Statistics
 Center project name: AAF
 Sequencing vector: pUC18, 108752
 Chemistry: Dye-terminator Big Dye, 100%
 *Consensus quality: 247096 at least Q20
 *Consensus quality: 242518 at least Q30
 *Consensus quality: 235589 at least Q40
 *Estimated insert size: agarose-FP - N/A
 *Estimated insert size: 254133 - sum-of-contigs
 Quality coverage: agarose-FP - N/A
 Quality coverage: 6 x in Q20 bases; sum-of-contigs estimation

 * NOTE: This is a 'working draft' sequence. It currently

* consists of 44 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 27415: contig of 27415 bp in length
* 27416 27435: gap of unknown length
* 27436 57817: contig of 30382 bp in length
* 57818 57837: gap of unknown length
* 57838 81637: contig of 23800 bp in length
* 81638 81657: gap of unknown length
* 81658 102155: contig of 20498 bp in length
* 102156 102175: gap of unknown length
* 102176 121200: contig of 19025 bp in length
* 121201 121220: gap of unknown length
* 121221 135073: contig of 13853 bp in length
* 135074 135093: gap of unknown length
* 135094 156689: contig of 21596 bp in length
* 156690 156709: gap of unknown length
* 156710 16185: contig of 9476 bp in length
* 16186 166205: gap of unknown length
* 166206 178798: contig of 12593 bp in length
* 178799 178818: gap of unknown length
* 178819 187247: contig of 8429 bp in length
* 187248 187267: gap of unknown length
* 187268 192492: contig of 5225 bp in length
* 192493 192512: gap of unknown length
* 192513 200084: contig of 7572 bp in length
* 200085 200104: gap of unknown length
* 200105 207054: contig of 6950 bp in length
* 207055 207074: gap of unknown length
* 207075 212944: contig of 5870 bp in length
* 212945 212964: gap of unknown length
* 212965 219468: contig of 6504 bp in length
* 219469 219488: gap of unknown length
* 219489 220369: contig of 881 bp in length
* 220370 220389: gap of unknown length
* 220390 221164: contig of 775 bp in length
* 221165 221184: gap of unknown length
* 221185 222796: contig of 1612 bp in length
* 222797 222816: gap of unknown length
* 222817 224762: contig of 1946 bp in length
* 224763 224782: gap of unknown length
* 224783 225999: contig of 1217 bp in length
* 226000 226019: gap of unknown length
* 226020 227409: contig of 1390 bp in length
* 227410 227429: gap of unknown length
* 227430 230145: contig of 2716 bp in length
* 230146 230165: gap of unknown length
* 230166 231842: contig of 1677 bp in length
* 231843 231862: gap of unknown length
* 231863 232786: contig of 924 bp in length
* 232787 232806: gap of unknown length
* 232807 233137: contig of 331 bp in length
* 233138 233157: gap of unknown length
* 233158 234519: contig of 1362 bp in length
* 234520 234539: gap of unknown length
* 234540 236416: contig of 1877 bp in length
* 236417 236436: gap of unknown length
* 236437 237967: contig of 1531 bp in length
* 237968 237987: gap of unknown length
* 237988 239220: contig of 1233 bp in length
* 239221 239240: gap of unknown length
* 239241 239367: contig of 127 bp in length
* 239368 239387: gap of unknown length
* 239388 239959: contig of 572 bp in length
* 239960 239979: gap of unknown length
* 239980 241016: contig of 1037 bp in length
* 241017 241036: gap of unknown length
* 241037 242391: contig of 1355 bp in length
* 242392 242411: gap of unknown length

FEATURES
source

* 242412 243823: contig of 1412 bp in length
* 243824 243843: gap of unknown length
* 243844 245227: contig of 1384 bp in length
* 245228 245247: gap of unknown length
* 245248 246334: contig of 1087 bp in length
* 246335 246354: gap of unknown length
* 246355 247451: contig of 1097 bp in length
* 247452 247471: gap of unknown length
* 247472 248780: contig of 1309 bp in length
* 248781 248800: gap of unknown length
* 248801 250041: contig of 1241 bp in length
* 250042 250061: gap of unknown length
* 250062 251173: contig of 1112 bp in length
* 251174 251193: gap of unknown length
* 251194 252713: contig of 1520 bp in length
* 252714 252733: gap of unknown length
* 252734 254014: contig of 1281 bp in length
* 254015 254034: gap of unknown length
* 254035 254176: contig of 142 bp in length
* 254177 254196: gap of unknown length
* 254197 254993: contig of 797 bp in length.
Location/Qualifiers
1. 254993
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL6/J"
/db_xref="taxon:10090"
/chromosome="6"
/clone="RP23-117123"
/sex="male"
1. 27415
/note="assembly_name:Contig138"
27436..57817
/note="assembly_name:Contig137"
clone_end:586
vector_side:left"
57838..81637
/note="assembly_name:Contig136"
81658..102155
/note="assembly_name:Contig135"
102176..121200
/note="assembly_name:Contig134"
121221..135073
/note="assembly_name:Contig133"
135094..156689
/note="assembly_name:Contig132"
156710..166185
/note="assembly_name:Contig131"
clone_end:177
vector_side:left"
166206..178798
/note="assembly_name:Contig130"
178819..187247
/note="assembly_name:Contig129"
187268..192492
/note="assembly_name:Contig128"
192513..200084
/note="assembly_name:Contig127"
200105..207054
/note="assembly_name:Contig126"
207075..212944
/note="assembly_name:Contig125"
212965..219468
/note="assembly_name:Contig124"
219469..220369
/note="assembly_name:Contig123"
220390..221164
/note="assembly_name:Contig122"
221185..222796
/note="assembly_name:Contig121"
222817..224762
/note="assembly_name:Contig120"
224783..225999

```

/misc_feature      /note="assembly_name:Contig119"
                   22620..227409
/misc_feature      /note="assembly_name:Contig118"
                   227430..230145
/misc_feature      /note="assembly_name:Contig117"
                   230166..231842
/misc_feature      /note="assembly_name:Contig116"
                   231863..232786
/misc_feature      /note="assembly_name:Contig115"
                   232807..233137
/misc_feature      /note="assembly_name:Contig114"

```

Alignment Scores:

```

Pred. No.:      3.2e-103      Length:      254993
Score:          1212.00      Matches:      230
Percent Similarity: 97.58%      Conservative: 12
Best Local Similarity: 92.74%      Mismatches: 6
Query Match:      83.59%      Indels:      0
DB:              2          Gaps:      0

```

US-09-896-522-2 (1-277) x AC078885 (1-254993)

```

Qy      30 GlyGlyThrAlaSerGlyLysSerThrValCysGluLysIleMetGluLeuGlyGln 49
Db      156682 GGCAGGACCCGCTAGTGGCAAGTCAACAGTGTGAGAAAGATCATGAGAGCTGTGGACAG 156623

Qy      50 AsnGluValGluGlnArgGlnArgLysValIleLeuSerGlnAspArgPheThrLys 69
Db      156622 AACCAAGGAGACCCGCGGACGCGAAGTGGTCACTCCGAGCCAGAGTGGCTTCTACAG 156563

Qy      70 ValLeuThrAlaGluGlnLysAlaLysAlaLeuLysGlyGlnTyraAsnPheAspHisPro 89
Db      156562 GTTCTGACGGCGGACGAGAAAGCCAGCTTTGAAGGACAGTCAATTTTGACACCA 156503

Qy      90 AspAlaPheAspAsnAspLeuMetHisArgThrLeuLysAsnIleValGluGlyLysThr 109
Db      156502 GATGCTTTGTGATATGATCTGATCACAGACCTCGTAAACATTTGTGAAGCGAAACT 156443

Qy      110 ValGluValProThrTyraAspPheValThrHisSerArgLeuProGluThrThrValVal 129
Db      156442 GTGAGGTCCTACCTAGTATTTGTGACCACTCAAGTTTACAGAGACCTGTGTC 156383

Qy      130 TyrProAlaAspValValLeuPheGluGlyIleLeuValPheTyrsSerGlnGluIleArg 149
Db      156382 TACCACGCTGATGTGTTCTGTCGAGGCACTTTGGTATCTACACCCGGAGATCCGG 156323

Qy      150 AspMetPheHisLeuAlaGluPheValAspThrAspSerAspValAlaGluSerArgArg 169
Db      156322 GACATGTTCCACCTGCGCTCTTGTGTGACACAGACTCTGATTTAGGCTGTCTCGAAGA 156263

Qy      170 ValLeuArgAspValArgArgGlyArgAspLeuGluGlnIleLeuThrGlnTyThrThr 189
Db      156262 GTTCTCGGAGTGTGCAACGAGAGGAGCTTGAACAGATCTGACTCAGTATACCGCC 156203

Qy      190 PheValLysProAlaPheGluGluPheCysLeuProThrLysTyraAlaAspValIle 209
Db      156202 TTTGTGAACACGAGCTTTGAGGAGTTCGCTGCGACTAAGAGTACGCTGAGTATC 156143

Qy      210 IleProArgGlyValAspAsnMetValAlaIleAsnLeuIleValGlnHisIleGlnAsp 229
Db      156142 ATCCCTCGAGAGTGTGATTAATGTGTGCGCATCACTGATCCGCAACATCCAGAC 156083

Qy      230 IleLeuAsnGlyAspIleCysLysValPheHisArgGlyGlySerAsnGlyArgSerTyLys 249
Db      156082 ATCTCTCAACGGGAGCTGTGACAGACGAGGAGCGGCGCCAAACGGCGCAACAG 156023

Qy      250 ArgThrPheSerGluProGlyAspHisProGlyMetLeuThrSerGlyLysArgSerHis 269
Db      156022 AGGACCTTCCCGGAGCAGAGATCAACCTGGGGTGTGGCCATGGCAACGCTCACAC 155963

Qy      270 LeuGluSerSerArgProHis 277
Db      155962 CTGGAGTCTAGCAGACAGACCCCAT 155939

```

```

RESULT 13
MUSURKIA          952 bp      DNA      linear      ROD 05-MAR-2001
LOCUS             Mus musculus uridine kinase pseudogene.
DEFINITION        L31784
ACCESSION         L31784.1 GI:471982
VERSION           pseudogene: uridine kinase.
KEYWORDS          Mus musculus (house mouse)
SOURCE            Mus musculus
ORGANISM          Mus musculus
REFERENCE          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS           Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE             1 (bases 1 to 952)
FEATURES           The functions and consensus motifs of nine types of peptide
                   segments that form different types of nucleotide-binding sites
JOURNAL            Eur. J. Biochem. 222 (1), 9-19 (1994)
MEDLINE            94259063
PUBMED            8200357
REFERENCE          2 (bases 1 to 952)
AUTHORS           Rapp, P.A. and Traut, T.W.
TITLE             Cloning and expression of a cDNA encoding uridine kinase from mouse
JOURNAL            Arch. Biochem. Biophys. 336 (1), 105-112 (1996)
MEDLINE            97108719
PUBMED            8951040
FEATURES           Location/Qualifiers
                   source
                   1..952
                   /organism="Mus musculus"
                   /mol_type="genomic DNA"
                   /db_xref="taxon:10090"
                   /tissue_type="kidney"
                   76..913
                   /standard_name="uridine-cytidine kinase"
                   /pseudo
                   /codon_start=1
                   /product="uridine kinase"
                   misc_difference 130
                   BASE COUNT      230 a      231 c      277 g      214 t
                   ORIGIN
                   Alignment Scores:
                   Pred. No.:      2.52e-101      Length:      952
                   Score:          1159.50      Matches:      234
                   Percent Similarity: 88.17%      Conservative: 12
                   Best Local Similarity: 83.87%      Mismatches: 31
                   Query Match:      79.97%      Indels:      2
                   DB:              10          Gaps:      1
                   US-09-896-522-2 (1-277) x MUSURKIA (1-952)
Qy      1 MetAlaSerAlaGlyGlyLysAspCysGluSerProAla---ProGluAlaAspArgPr 19
Db      76 ATGGCTTCGGCGGAGGCGGTGCTCAGAGAGTACGCGCTGCGAGGCGGATCTTCC 135

Qy      19 ohSglnArgProPheLeuIleGlyValSerGlyGlyThrAlaSerGlyLysSerThrVa 39
Db      136 TCAGCTTGCGCTTCTCATCGGCTGTGAGTGGCGGCGCCGCTAAGTGGCAAGTCAAG 195

Qy      39 lCysGluLysIleMetGluLeuLeuGlyGlnAsnGluValGluGlnArgGlnArgLysVa 59
Db      196 GTGTGAGAGATATGAGAGCTGCTGGGACAGATGAAGTGAAGTGAAGTGAAGTGAAGTGA 255

Qy      59 lValIleLeuSerGlnAspArgPheTyrlsValLeuThrAlaGluGlnLysAlaLysAl 79
Db      256 GGTCAATCTGAGCGACGAGCTGTTTACAAAGTTTGAAGGCTGAGACGAGAAAGCCAAAGC 315

Qy      79 AleuLysGlyGlnTyraAsnPheAspHisProAspAlaPheAspAsnAspLeuMetHisAr 99
Db      316 TTTGAAGGACAGTACAAATTTTGACCATGACAGATGCTTTGTGATTAATGATCTGATGACA 375

Qy      99 gThrLeuLysAsnIleValGluGlyLysThrValGluValProThrTyraAspPheValTh 119

```



```

Db      376 GACCTTAAAGCAATTTTGAAGCAAGTTGTTAAGCTTCATACCATGATTTTGGAC 435
Qy      119 rHiseraTgLeuProGluThrValValTyrProAlaAspValValLeuPheGluG1 139
Db      436 CGACTCAAGGTTTCCAGTGACCACTGTGTCTACACATGATTCGGCTTCTGTTCCAGAG 495
Qy      139 ylleleValPheTyrSerGlnGlnIleAraGspMecPheH1sleuAraGleuPheValAe 159
Db      496 CATCTTGATCTTACACTCAAGAGATCCGAGACATTTTCACTGGCGCTTTTGTGGA 555
Qy      159 pTHrAspSerAspValArgLeuSerArgValLeuAraGspValAraGspValArgG1ArgAs 179
Db      556 CACAGACCCTGATGTTAGCTGTCTTGAAGAGTTCTCCGGGATGTGCAAGAGGAAGCA 615
Qy      179 pLeuGluGlnIleLeuThrGlnTyrThrPheValLysProAlaPheGluGluPheCy 199
Db      616 CCGTGAACAGATCTGACTCAATACCGCCCTTGTGTAACACAGCCCTTGAGGAGTTCTG 675
Qy      199 sleuProThrLysLysTyrValAlaAspValIleleProArgLysValAlaAspMecValAl 219
Db      676 CCGGCCACTTGAAGAGTACCTGACGTGATCATCTCCGAGGGCTTGAACAATATGAGTGC 735
Qy      219 aileAsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysLysTrpH1 239
Db      736 CATCAACCTGATCGTATTAACACATCCAGACATCTCTCAACGGGAGCTGTGCAAGCGCA 795
Qy      239 sATGGLYGLYSerAsnGlyArgSerTyrLysArgThrPheSerGluProGlyAspHisPr 259
Db      796 CCGAGGGGGGGCCCAATGGGCGCAGCCCAAGAGGACCTTACCCGAGTCAGAGATCAACC 855
Qy      259 ogLMeLeuThrSerGlyLysArgSerHisleuGluSerSerSerTrpHis 277
Db      856 TGGGGTGTGCCACTGGTAAACGCTGACACCTGGAGTCTTACGACGACGCCAT 910

```

RESULT 14
AC118474
LOCUS
DEFINITION Mus musculus clone RP23-288012, *** SEQUENCING IN PROGRESS ***, 4
unoriented pieces.

AC118474
AC118474.4 GI:31074905
HTG: HTGS PHASE1; HTGS FULLTOP; HTGS_ACTIVEFIN.
MUS musculus (house mouse)
MUS musculus

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
MUS musculus
Birren, B., Nussbaum, C. and Lander, E.
1 (bases 1 to 192336)
2 (bases 1 to 192336)
Unpublished

2 (bases 1 to 192336)
Birren, B., Linton, L., Nussbaum, C., Lander, E., All, A., Allen, N.,
Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L.,
Boukigallier, B., Brown, A., Camarata, J., Campopiano, A., Chang, J.,
Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A.,
Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S., Dodge, S.,
Faro, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S.,
Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-pierre, N.,
Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kells, C., Labocque, K., Lamazares, R.,
Landers, T., Lehoczy, J., Levine, R., Lindblad-Toh, K., Liu, G.,
Maclean, C., Macdonald, P., Major, J., Matthews, C., McCarthy, M.,
McEwan, P., McKernan, K., Meldrum, J., Matthews, C.,
Mhova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R.,
Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D.,
Oliver, J., Peterson, K., Phunkhang, P., Plerre, N., Pollara, V.,
Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupack, R.,
Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Straus, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J.,
Topham, K., Travers, M., Travis, N., Triggillo, J., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G.,

TITLE
JOURNAL
REFERENCE
AUTHORS
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
Submitted (17-APR-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 192336)
Birren, B., Nussbaum, C., Lander, E., Abouelleil, A., Allen, N.,
Anderson, S., Arachchi, H. M., Barna, N., Bastien, V., Bloom, T.,
Boguslavsky, L., Boukigallier, B., Camarata, J., Chang, J., Choepel, Y.,
Collymore, A., Cook, A., Cooke, P., Corum, B., Dearellano, K.,
Diaz, J. S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S.,
Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S.,
Graham, L., Grand-pierre, N., Hafez, N., Hagopian, D., Hagos, B.,
Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R.,
Lindblad-Toh, K., Liu, G., Lui, A., Mabbitt, R., Maclean, C.,
Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M.,
Meldrum, J., Meneus, L., Mhova, T., Mlenga, V., Murphy, T., Naylor, J.,
Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P.,
O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Plerre, N.,
Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P.,
Roman, J., Schauer, S., Schupack, R., Seaman, S., Severy, P., Smith, C.,
Spencer, B., Stange-Thomann, N., Stojanovic, N., Strubs, M.,
Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M.,
Vassiliev, H., Venkataraman, V. S., Viel, R., Vo, A., Wilson, B., Wu, X.,
Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE
JOURNAL
COMMENT
Submitted (25-MAY-2003) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On May 25, 2003 this sequence version replaced g1:28975887.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence.submissions@genome.wi.mit.edu
----- Project Information
Center project name: L23680
Center clone name: 288_O12

***** NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 36631: contig of 36631 bp in length
* 36632 36731: gap of 100 bp
* 36732 160664: contig of 123933 bp in length
* 160665 160764: gap of 100 bp
* 160765 177955: contig of 17191 bp in length
* 177956 178055: gap of 100 bp
* 178056 192336: contig of 14281 bp in length.

FEATURES
source
1. 192336
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone="RP23-288012"
/clone_1lb="RP23-288012" Female Mouse BAC"

BASE COUNT 58391 a 36846 c 37829 g 58262 t 1008 others
ORIGIN

Alignment Scores:
Pred. No.: 2 1e-90 Length: 192336
Score: 1076.00 Matches: 225
Percent Similarity: 85.304 Conservative: 13
Best Local Similarity: 80.654 Mismatches: 36
Query Match: 74.214 Indels: 6
DB: 2 Gaps: 2

```

US-09-896-522-2 (1-277) x AC118474 (1-192336)
Qy      1 MetAlaserAlaGlyGlyGluAspCysGluSerProAla-----ProGluAlaAspArgPr 19
Db      77706 ATGGCTTGGCGGAGGCGGTGGCTCAAGAGTACTGGCGCTGCCGAGGCGCATCTTCC 77765
Qy      19 ohIsgInArpProPheLeuIleGlyValSerGlyGlyThrAlaSerGlyLysSerThrVa 39
Db      77766 TCAGCCTTGGCGGCGTCTCATCGGCTGAGTGGCGGCACTGATGGCAAGTCAACAGT 77825
Qy      39 lCysGluLysIleMetCgluLeuLeuGlyGlnAsnGluValGluGlnArgGlnArgLysVa 59
Db      77826 GTGTGAGAGAGATCATGAGCTGCTGGGACAAATGAACTGACCTCCGCTGAGCCCAAGTT 77885
Qy      59 lValIleLeuSerGlnAspArgPheThrLysValLeuThrAlaGluGlnLysAlaLysAl 79
Db      77886 GGTCATCTTGAGCCAGGAGCTGCTTCTACAGGCTTCTGAGGCTGAGCAAGAAAGCCAAAGGC 77945
Qy      79 aLeuLysGlyGlnTyraAsnPhaAspHisProAspAlaPheAspAsnAspLeuMetHisAr 99
Db      77946 TTGAGAGGACAGATACATTTTGACCATGAGATGCTTTGATATGATCTGATGAGACA 78005
Qy      99 gThrLeuLysAsnIleValGluGlyLysThrValGluValProThrTyraAspPheValTh 119
Db      78006 GACCTTGAAAAACATTTATGAAAGGCAAGTTGTTAGGTTCTTACCTATGATTTTGTGAC 78065
Qy      119 rHisSerArgLeuProGluThrThrThrValValTyraProAlaAspValValLeuPheGlu 139
Db      78066 CCACCTACAGTTTACCACTGACGACCACTGGTCTTACCAATGATGCGGTTCTGTTGAGGG 78125
Qy      139 yIleLeuValPheTySerGlnGluIleArgAspMetPheHisLeuArgLeuPheValAs 159
Db      78126 CATCTTGATTTCTTCACTCAAGAGATCCGGGACATGTTTCACTGCGCCTTTTGTGGA 78185
Qy      159 pThrAspSerAspValArgLeuSerArgArgValLeuArgAspValArgArgGlyArgAs 179
Db      78186 CACAGACCCCTGATGTAGGCTGCTTGTGAAGAGTTCTGCGGATGTGCAATGAGAGAGGGA 78245
Qy      179 pLeuGlnGlnIleLeuThrGlnIleTyraThrPheValLysProAlaPheGluGluPheCy 199
Db      78246 CTTGAGACAGATCTTCACTGCTGACGACCACTGTTGTAATGCGCTTGTGAGAGTTCTG 78305
Qy      199 sLeuProThrLysLysTyraAlaAspValIleIleProArgGlyValAspAsnMetVal 219
Db      78306 CCTGCCGACCTTA-----GAAAGTATCATCTCCCTGAGGCTTGACATATGCTGGC 78355
Qy      219 aIleAsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysLysThrPh 239
Db      78356 CATCAACCTGATCGTATTAACATCCAGACATCTCTCAACCGGAGCCTGTGCAAGCGGCA 78415
Qy      239 sArgGlyLysSerAsnGlyArgSerTyraArgThrPheSerGluProGlyAspPheIle 259
Db      78416 CCGAGAGAGGCCCAATGGGACAGCCACACAGAGAGACCTTACCAGATCAGAGAGATCACCC 78475
Qy      259 oGlyMetLeuThrSerGlyLysArgSerHisLeuGlnSerSerSerArgProHis 277
Db      78476 TGGGGTGTGGCCACTGGTAAACCTGACACTGGAAGTCTAGCAGCAGATCCCAT 78530

RESULT 15
AC076974      211075 bp      DNA      linear      ROD 16-NOV-2002
LOCUS      Mus musculus chromosome 8 clone rp23-267j18 strain C57BL/6J,
DEFINITION      complete sequence.
ACCESSION      AC076974
VERSION      AC076974
KEYWORDS      GI:25046382
SOURCE      Mus musculus (house mouse)
ORGANISM      Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE      1 (bases 1 to 211075)
AUTHORS      Deschamps,S., Gu,W. and Roe,B.A.

```

```

TITLE      Mus musculus Chromosome 1 BAC Clone rp23-267j18
JOURNAL      Unpublished
REFERENCE      2 (bases 1 to 211075)
AUTHORS      Deschamps,S., Gu,W. and Roe,B.A.
TITLE      Direct Submission
JOURNAL      Submitted (01-AUG-2000) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
REFERENCE      3 (bases 1 to 211075)
AUTHORS      Deschamps,S., Be, Ziemmer,B., Gu,W. and Roe,B.A.
TITLE      Direct Submission
JOURNAL      Submitted (12-JUL-2002) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
REFERENCE      4 (bases 1 to 211075)
AUTHORS      Deschamps,S., Be, Ziemmer,B., Gu,W. and Roe,B.A.
TITLE      Direct Submission
JOURNAL      Submitted (18-JUL-2002) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
REFERENCE      5 (bases 1 to 211075)
AUTHORS      Deschamps,S., Be, Ziemmer,B., Gu,W. and Roe,B.A.
TITLE      Direct Submission
JOURNAL      Submitted (16-NOV-2002) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
COMMENT      On Nov 16, 2002 this sequence version replaced gi:21734996.
----- Genome Center
Center: Department Of Chemistry And Biochemistry
The University Of Oklahoma
Center code:UOKNOR
-----
FEATURES
source
1..211075
location/Qualifiers
1..211075
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/chromosome="8"
/clone="rp23-267j18"
/clone_lib="RPCT - 23 Female (C57BL/6J) Mouse BAC Library"
/notes="This is one of two clones in the same well from
rp23-267j18"
BASE COUNT      59757 a 44080 c 44532 g 62706 t
ORIGIN
Alignment Scores:
Pred. No.:      2,376-90      Length:      211075
Score:      1076.00      Matches:      225
Percent Similarity:      85.30%      Conservative:      13
Best Local Similarity:      80.65%      Mismatches:      36
Query Match:      74,21%      Indels:      6
DB:      10      Gaps:      2
US-09-896-522-2 (1-277) x AC076974 (1-211075)
Qy      1 MetAlaserAlaGlyGlyGluAspCysGluSerProAla-----ProGluAlaAspArgPr 19
Db      195148 ATGGCTTGGCGGAGGCGGTGGCTCAAGAGTACTGGCGCTGCCGAGGCGCATCTTCC 195207
Qy      19 ohIsgInArpProPheLeuIleGlyValSerGlyGlyThrAlaSerGlyLysSerThrVa 39
Db      195208 TCAGCCTTGGCGGCGTCTCATCGGCTGAGTGGCGGCACTGATGGCAAGTCAACAGT 195267
Qy      39 lCysGluLysIleMetCgluLeuLeuGlyGlnAsnGluValGluGlnArgGlnArgLysVa 59
Db      195268 GTGTGAGAGAGATCATGAGCTGCTGGGACAAATGAACTGACCTCCGCTGAGCCCAAGTT 195327
Qy      59 lValIleLeuSerGlnAspArgPheThrLysValLeuThrAlaGluGlnLysAlaLysAl 79
Db      195328 GGTCATCTTGAGCCAGGAGCTGCTTCTACAGGCTTCTGAGGCTGAGCAAGAAAGCCAAAGG 195387
Qy      79 aLeuLysGlyGlnTyraAsnPhaAspHisProAspAlaPheAspAsnAspLeuMetHisAr 99

```

```
Db 195388 TTGAAAGGACAGACATTTTGCATGCAGATGCTTTGATATGATCGATGACAA 195447
QY 99 gThrLeuLysAsnLLeValGluGlyLysThrValGluValProThrTyrAspPheValTh 119
Db 195448 GACCTTGAAAACATTATGTAAGGCMAAGTGTATAGGTTCTTACCTATGATTTTGTGAC 195507
QY 119 rHisSerArgLeuProGluThrThrValValTyrProAlaAspValValLeuPheGlu 139
Db 195508 CCACTCAAGGTTACCAAGTACCACTGTGCTTACCAATGATGCGGTTCTGTTGAGGG 195567
QY 139 YIleLeuValPheTyrSerGluGluLeuArgAspMetPheHisLeuArgLeuPheValAs 159
Db 195568 CATCTTGATTTCTACACTCAAGAGATCCGGGACATGTTTCACTGCGCTTGTGGA 195627
QY 159 pThrAspSerAspValArgLeuSerArgValLeuArgAspValArgArgGlyArgAs 179
Db 195628 CACAGACCCTGATGTGTAGGCTGTGTAAGAGTTCTCTGGGATGTGCAATGAGGAAAGGA 195687
QY 179 pLeuGluGluLLeuThrGlnTyrThrThrPheValLysProAlaPheGluGluPheCy 199
Db 195688 CTTGAGCAATCTTCTGACTGACGACCACTTGTGAAATCGGCTTCAAGAGTCTG 195747
QY 199 sLeuProThrLysLysTyrAlaAspValLLeuLeuProArgGlyValAspAsnMetValAl 219
Db 195748 CTGCGCACTTA-----GAAGTATCATTCCTCGAGGGCTTGACAAATATGTGCGC 195797
QY 219 aIleAsnLeuLLeuValGlnHisIleGlnAspLLeuAsnGlyAspLLeuCyLysTPrH 239
Db 195798 CATCAACCTGATCTATACACATCCAGACATCTCAACGGGACCTGTGCAAGCGGCA 195857
QY 239 sArgGlyGlySerAsnGlyArgSerTyrLysArgThrPheSerGluProGlyAspHisPr 259
Db 195858 CCGAGAGGGGCCCAATGGGCAACAGCCACAAAGAGACCTTAACCGAGTCAGAGATCACCC 195917
QY 259 oGlyMetLeuThrSerGlyLysArgSerHisLeuGluSerSerSerArgProHis 277
Db 195918 TGGGGTGTGGCCACTGTAAACGCTGACACCTGAGTCTAGCAGCAGATCCCAT 195972
```

Search completed: November 25, 2003, 09:29:17
Job time : 4279 secs

This Page Blank (uspto)


```

QY 89 CCGAGATGCTTGGCGGAGGCGAAGACTGCGAGACCCCGCGCGGAGCCGACCGTC 148
D 90 CCGAGATGCTTGGCGGAGGCGAAGACTGCGAGACCCCGCGCGGAGCCGACCGTC 149
QY 149 CGCAGCAGCGGCGCTTCTGATAGAGGGTGAAGCGGCGGCACTGCGAGGGGAAGTCAGCG 208
D 150 CGCAGCAGCGGCGCTTCTGATAGAGGGTGAAGCGGCGGCACTGCGAGGGGAAGTCAGCG 209
QY 209 TGTGTGAAGAGATCATGAGAGTTGCTGAGACAGAAAGAGTGAAGACAGCGGCGAGAGG 268
D 210 TGTGTGAAGAGATCATGAGAGTTGCTGAGACAGAAAGAGTGAAGACAGCGGCGAGAGG 269
QY 269 TGTGTGAAGAGATCATGAGAGTTGCTGAGACAGAAAGAGTGAAGACAGCGGCGAGAGG 328
D 270 TGTGTGAAGAGATCATGAGAGTTGCTGAGACAGAAAGAGTGAAGACAGCGGCGAGAGG 329
QY 329 CTTGGAAGAGACATGATATTTTGAACCATCCAGATGCTTGTATATGATTTGATGACA 388
D 330 CTTGGAAGAGACATGATATTTTGAACCATCCAGATGCTTGTATATGATTTGATGACA 389
QY 389 GGAATCTGAAGACATGATGAGGCGAAGAGGAGTGAAGTGAAGTGAAGTGAAGTGAAG 448
D 390 GGAATCTGAAGACATGATGAGGCGAAGAGGAGTGAAGTGAAGTGAAGTGAAGTGAAG 449
QY 449 CACACTGAAGGTTACAGAGACAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 508
D 450 CACACTGAAGGTTACAGAGACAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 509
QY 509 GCAATCTGTGTTCTTACAGAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 568
D 510 GCAATCTGTGTTCTTACAGAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 569
QY 569 ACACCGAGCTCCGAGCTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 628
D 570 ACACCGAGCTCCGAGCTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 629
QY 629 ACCTGAGGAGATCTGAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 688
D 630 ACCTGAGGAGATCTGAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 689
QY 689 GCCTGCGGAGCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 748
D 690 GCCTGCGGAGCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 749
QY 749 CCATCAACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 808
D 750 CCATCAACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 809
QY 809 ACCGAGAGGAGTCCATGAGGAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 868
D 810 ACCGAGAGGAGTCCATGAGGAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 869
QY 869 CTGGAGATGCTGACCTTGGCAAGCGGTCAATTTTGAATGATGATGATGATGATGATGATG 928
D 870 CTGGAGATGCTGACCTTGGCAAGCGGTCAATTTTGAATGATGATGATGATGATGATGATG 929
QY 929 GGGTCTGCGAGCTTCAAGGAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 988
D 930 GGGTCTGCGAGCTTCAAGGAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 989
QY 989 GGGAGCGGACCCGACACCACTGCT 1013
D 990 GGGAGCGGACCCGACACCACTGCT 1014

```

```

RESULT 2
US-09-536-647-1
; Sequence 1, Application US/09536647
; Patent No. 6579708
; GENERAL INFORMATION
; APPLICANT: Ho, Yen Sen
; APPLICANT: Johnson, Randall
; TITLE OF INVENTION: No. 6579708e1 Human Uridine Kinase

```

```

; FILE REFERENCE: GP50020
; CURRENT APPLICATION NUMBER: US/09/536,647
; CURRENT FILING DATE: 2000-03-27
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 834
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(780)
US-09-536-647-1

Query Match 51.3%; Score 832.4; DB 4; Length 834;
Best Local Similarity 99.9%; Pred. No. 1.8e-212;
Matches 833; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 94 ATGGCTTCGCGGAGGCGAAGACTGCGAGAGCCCGCGGAGGCGGACCGTCCGAC 153
D 1 ATGGCTTCGCGGAGGCGAAGACTGCGAGAGCCCGCGGAGGCGGACCGTCCGAC 60
QY 154 CAGCGGCGCTTCTGATAGAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 213
D 61 CAGCGGCGCTTCTGATAGAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 120
QY 214 GAGAAATCATGAGAGTGTGCTGAGACAGAAAGAGTGAAGTGAAGTGAAGTGAAGTGAAG 273
D 121 GAGAAATCATGAGAGTGTGCTGAGACAGAAAGAGTGAAGTGAAGTGAAGTGAAGTGAAG 180
QY 274 ATCTGAGCAGAGCAGGTTCTTCAAGGCTCTGAGCGGAGAGAGCAAGGCGGAGCTTG 333
D 181 ATCTGAGCAGAGCAGGTTCTTCAAGGCTCTGAGCGGAGAGAGCAAGGCGGAGCTTG 240
QY 334 AAGAGCAGTACATTTTGAACCATCCAGATGCTTGTATGATTTGATGACAGAGCT 393
D 241 AAGAGCAGTACATTTTGAACCATCCAGATGCTTGTATGATTTGATGACAGAGCT 300
QY 394 CTGAAGACATGATGAGGAGGCAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 453
D 301 CTGAAGACATGATGAGGAGGCAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 360
QY 454 TCAAGGTTACAGAGCAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 513
D 361 TCAAGGTTACAGAGCAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 420
QY 514 TTGATGTTCAAGCAGAGAGATCCGAGATGATGATGATGATGATGATGATGATGATGATG 573
D 421 TTGATGTTCAAGCAGAGAGATCCGAGATGATGATGATGATGATGATGATGATGATGATG 480
QY 574 GACTCCGAGCTGAGTGTCTGAGAGATTTCCGAGAGTGAAGTGAAGTGAAGTGAAGTGAAG 633
D 481 GACTCCGAGCTGAGTGTCTGAGAGATTTCCGAGAGTGAAGTGAAGTGAAGTGAAGTGAAG 540
QY 634 GAGCAGATCTGAGCAGATGACACCACTTGTGAAGCCGAGCTTCAAGAGATTTCTGCTG 693
D 541 GAGCAGATCTGAGCAGATGACACCACTTGTGAAGCCGAGCTTCAAGAGATTTCTGCTG 600
QY 694 CGCAGAAAGAGTATCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 753
D 601 CGCAGAAAGAGTATCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 660
QY 754 AACCTGATGATGAGACATCCAGGACATTTGAAATGATGATGATGATGATGATGATGATG 813
D 661 AACCTGATGATGAGACATCCAGGACATTTGAAATGATGATGATGATGATGATGATGATG 720
QY 814 GAGAGGTCATAGGAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 873
D 721 GAGAGGTCATAGGAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 780
QY 874 ATGCTGACCTTGGCAAGCGGTCACTTTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 927
D 781 ATGCTGACCTTGGCAAGCGGTCACTTTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 834

```

RESULT 3
US-09-221-017B-368
Sequence 368, Application US/09221017B
Patent No. 6444799
GENERAL INFORMATION:
APPLICANT: Roess, Bruce C.
TITLE OF INVENTION: P. GINGIVALIS NUCLEOTIDES AND USES THEREOF
NUMBER OF SEQUENCES: 1120
CORRESPONDENCE ADDRESSES:
ADDRESSEE: MORRISON & FOERSTER
STREET: 755 PAGE MILL ROAD
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304-1018
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows
SOFTWARE: FASTSEQ for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/221,017B
FILING DATE: 23-DEC-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: P1182
FILING DATE: 31-DEC-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: P1546
FILING DATE: 30-JAN-1998
PRIOR APPLICATION DATA: P22911
FILING DATE: 09-APR-1998
APPLICATION NUMBER: PCT/AU98/01023
FILING DATE: 10-DEC-1998
ATTORNEY/AGENT INFORMATION:
NAME: Montoy, Gladys H
REGISTRATION NUMBER: 32,430
REFERENCE/DOCKET NUMBER: 27340-20021.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-813-5600
TELEFAX: 650-494-0792
TELEX: 706141
INFORMATION FOR SEQ ID NO: 368:
SEQUENCE CHARACTERISTICS:
LENGTH: 5687 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: UNKNOWN
ORIGINAL SOURCE:
ORGANISM: PORPHYROMONAS GINGIVALIS
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1...5687
US-09-221-017B-368
Query Match 6.6%; Score 106.8; DB 4; Length 5687;
Best Local Similarity 54.3%; Pred. No. 1,6e-18;
Matches 238; Conservative 0; Mismatches 197; Indels 3; Gaps 1;
QY 341 AGTACATTTTGGACATCCAGATGCTTTGATATGATTGATGACAGGACTCTGAAGA 400
DB 4854 AGCTGAAGTACGACATCCATGCTTTTCACACGAGATATGTCGTACGAGATATTTCTCT 4913
QY 401 ACATCGTGAGGCGCAAAACGGTGAGGTGCGGACCTATGATTTTGTGACACACTCAAGGT 460
DB 4914 CTCTGAAGGAGGCAAAACGATAGAGGTCGGGTATATCTTTCTGTAAGACACATCGTT 4973

QY 461 TACAGAGACACAGGTGCTTACCCCTGCGAGCTGCTTCTGTTGAGGCAATCTGTGT 520
DB 4974 TCGAAGAAAAAGTACCGTTTCCTGCGCAAAAGTATCTGATGATGAAATCTGATAT 5033
QY 521 TCTACAGCAGAGATCCGGGACATGTTCCACCTGCGCTCTTCTGTGACACCGACTCG 580
DB 5034 TCGAAGAACAAAGAGCTGCGGAGATCTGATGATGTAAGATATTTCTGATACCGATGCG 5093
QY 581 ACCTCAGCTGTCTGGAAGTTCTCGGAGCT--GCGCGAGGAGGAGACTTGAGC 637
DB 5094 ATATTCGATTGGCGCGCCCTTGTCGCGATGTCAGGAGACGGGAGCCCAATATGAT 5153
QY 638 AGATTCTGACGAGTACACACCTTCGTGAAGCGGCGCTTCGAGGAGTCTGCTCCGA 697
DB 5154 CGGTATTTGACACATCTTACGATCGGCTGCTATGACAGAGATTTTGTGAAACAT 5213
QY 698 CAAGAAGATGCGCATGATGATCCACAGAGAGTGACATATGTTGCCATCAACC 757
DB 5214 CCAAGCGGTACCGCATCTGATCATTCGGAAGGTGATTCAATTCGTCGCTCTCAC 5273
QY 758 TGATCGTGAAGACATCC 775
DB 5274 TCCTTGTGCAAAAAATCC 5291
RESULT 4
US-08-961-527-76/c
Sequence 76, Application US/08961527
Patent No. 6420135
GENERAL INFORMATION:
APPLICANT: Charles Kunsch
TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
NUMBER OF SEQUENCES: 391
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
OPERATING SYSTEM: HP Vectra 486/33
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/961,527
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB340P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 76:
SEQUENCE CHARACTERISTICS:
LENGTH: 10011 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-961-527-76
Query Match 6.3%; Score 101.8; DB 4; Length 10011;
Best Local Similarity 53.2%; Pred. No. 4.5e-17;
Matches 239; Conservative 0; Mismatches 207; Indels 3; Gaps 1;
QY 345 CATTTTTACATCCAGATGCTTTGATATGATTTGATGACAGAGACTCTGAAGAACAT 404

```
Db 10001 CAACCTATGACCATCTTTTGGCTTTGATACGACTTGATGATCGACGACATTTAAGCAATT 9942
Qy 405 CGTGGAGGGCAAAAAGGTGAGAGTCCGACCTATGATTTTGTGACACACTCAAGTTACC 464
Db 9941 GTTGGACAGGGGCTCGGTGGACATCCGACCTACGATATACAGAGATACAGAGTAG 9882
Qy 465 AGAAGACCAAGGTGTCTACCTCGGACGTGTCTGTGTTAGGGCATTTGGTGTCTA 524
Db 9881 CAAGACCTATGTCGAGGAACCTCAAGATGCTTTATGTTGAGGGTATTTTGGCTTGA 9822
Qy 525 CAGCAGAGAGATCCGGGACATGTTCCACCTCGGCTCTGTCGAGACCCGACCTCGAG 584
Db 9821 GGAACAAGCTGTGGCGCATTTGATGATATCAAGATTTTGTGATACGATGCGATGT 9762
Qy 585 CAGGCTGTCTCGAAGAGTCTCCGGGA---CGTCCGCGAGGAGGAGGACCTGAGCAGAT 641
Db 9761 GCGCATTTATGTCGTATCAAGCGTATGAGAGAGGCGTGGCCGTAGCCTTGATAGCGT 9702
Qy 642 TCTGACGACATACACCACTTCGTGMAAGCCGCTTGAGAGAGTTTCTGCCCGACAA 701
Db 9701 TATTAAACGATGCTTAGGTGTGTCMAACCAATGATACCAAGTTTATCGAGTCACTAA 9642
Qy 702 GAAGTATGCCGATGATGATCCACGAGAGTGAACAATATGGTGCATCACTGAT 761
Db 9641 GCGTATGCTGATATCGTCAATCTCGAAGGGTTAGCAATACGTTGCTGACCTGTT 9582
Qy 762 CGTGCAGACATCCAGACATTTGTAATG 790
Db 9581 GACGACCAAGATTCGCAAAATTTTGAAG 9553
```

```
RESULT 5
US-09-134-001C-781
; Sequence 781, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 781
; LENGTH: 651
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-781
```

```
Query Match 5.3%; Score 86.2; DB 4; Length 651;
Best Local Similarity 51.5%; Pred. No. 18-13;
Matches 224; Conservative 0; Mismatches 208; Indels 3; Gaps 1;

Qy 346 AATTGACCATCGAGATGCTTTGATATGATTTGATGACAGACTCTGAGAACATC 405
Db 196 AATTATGACCACTTATGATTTGATATGATTTACTTATTCAATTAATTTAAAGATTTA 255
Qy 406 GTGAGGGCAAAAGGTGAGAGTCCGACCTATGATTTTGTGACACACTCAAGTTACA 465
Db 256 AGAATGAAAAACAGTAGAGTACTACGATCAATTAATCTCAACATACACGTAAGTAA 315
Qy 466 GAGACCAAGGTGTCTACCTCGGAGAGTGTCTGTTGAGGGACCTTGGGTGTCTAC 525
Db 316 GAAACAAATTGATTTGATTCAAAAGATGTTATTCGTAGAGGATCTTTGCTTGA 375
Qy 526 AGCCAGAGATCCGGGACATGTTCCACTGCGCTCTTGTGACACCGACTCCGAGCTC 585
Db 376 AAGAACACATTAAGAGACATGATGATGTAATAATTTATGTCGATACGATGCTGATTTA 435
```

```
Qy 586 AGCTGTCTGGAAGAGTTCTCCGGGA---CGTCCGCGAGGAGGAGGACCTGAGCAGATT 642
Db 436 CGAATTTTAAAGTGGGTTTACAAAGAGACTGTGTCGTAACAATGATCAGTA 495
Qy 643 CTGACGAGATACCAACCTTGTGGAAGCGGCTTCGAGAGAGTTCTGCCCTCGCAAAAG 702
Db 496 ATTATACATATCTTTAATGTAGTAGACCTATGACATATTTATTTGAACCTACAAA 555
Qy 703 AAGTATCCGATGTGATCATCCACGAGAGTGAACAATATGTTGCCATCAACTGATC 762
Db 556 AAGCATGTCGACATCAATATCTTGAAGAGGAGCAGCAATTAAGTTCCATATATATAG 615
Qy 763 GTGACGACATCCAG 777
Db 616 ACTACTAAATCCAG 630
```

```
RESULT 6
US-09-107-532A-1370
; Sequence 1370, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESS: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; City: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Arinello, Pamela Denek
; REGISTRATION NUMBER: 40,489
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 1370:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 633 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; MOLECULE TYPE: circular
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...633
; SEQUENCE DESCRIPTION: SEQ ID NO: 1370:
US-09-107-532A-1370

Query Match 4.5%; Score 73.4; DB 4; Length 633;
Best Local Similarity 50.1%; Pred. No. 4,6e-10;
```


	Matches	210, Conservative	0, Mismatches	206, Indels	3, Gaps	1,
Qy	346	AATTTGACCATTCAGATGCCCTTTGATTAATGATTGATGCAACAGACTTGAGAAATC	405			
Db	178	AATTACGATCCACCATTTGGCTTTGATATATATCTGCTGATTGACATGTAAGGACCTA	237			
Qy	406	GTGAGGGGCAAAACGTGAGAGTGCAGACTATGATTTTGTGACACTCAAGGTTTACA	465			
Db	238	CTGAACCTATAAGCATTTGAAAAACCGGTATATGATATATGTAAGGCATACAGAAAGTACG	297			
Qy	466	GAGACCAACGCTGTCTTACCTCGGACGCTGTTCTGTTTAGGCGCATCTTGTTCTAC	525			
Db	298	GCAACGATCATTTACAGAAACAAAGAAAGTATCATTTAGAAAGAAATCCTGATTTTGA	357			
Qy	526	AGCCAGAGATCCGGGACATGTTCCACTGCGCCTTTCTGTGACACCGACTCCGACGTC	585			
Db	358	GACCAACGCTCGCTGACTGATGATATCAAAAGTATCGTAGATACAGATGATGACATT	417			
Qy	586	AGGCTGTCTCGAAAGATTCTCCGGAGCGTG--CGCCGAGGAGAGGACCTTGAGACATT	642			
Db	418	CGTATCATTCCTCGTATCAAAACGGGATATGGAAGACGGGAGCCACACTTGCTCGTG	477			
Qy	643	CTGACGCAGTACACCACTTCTGTGAAGCCGCGCTTCAGAGAGTTCTGCTGCCGACAAG	702			
Db	478	ATCGAACATATTTTGACAGTGTGTTAAACCGATGTATCATCAATTTATGAAACCAACG	537			
Qy	703	AAGTATGCCATGTGATATATCCCAAGAGAGATGACAAATATGTGTGCATCAACTGAT	761			
Db	538	CGCATATGGGATATCATTTGTTCAGAAAGCCGAGAAACCAACGTTGCTATTTGATTTGAT	596			

```

RESULT 7
US-09-198-452A-1/c
: Sequence 1, Application US/09198452A
: Patent No. 6559294
: GENERAL INFORMATION:
: APPLICANT: Griffiths, R.
: TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
: TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention
: TITLE OF INVENTION: and treatment of infection
: FILE REFERENCE: 9710-003-999
: CURRENT APPLICATION NUMBER: US/09/198,452A
: CURRENT FILING DATE: 1998-11-24
: NUMBER OF SEQ ID NOS: 6849
: SEQ ID NO 1
: LENGTH: 1230025
: TYPE: DNA
: ORGANISM: Chlamydia pneumoniae
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (1)..(15000)
: OTHER INFORMATION: n=a or c or g or t
: NAME/KEY: misc_feature
: LOCATION: (15001)..(30000)
: OTHER INFORMATION: n=a or c or g or t
: NAME/KEY: misc_feature
: LOCATION: (30001)..(45000)
: OTHER INFORMATION: n=a or c or g or t
: NAME/KEY: misc_feature
: LOCATION: (45001)..(60000)
: OTHER INFORMATION: n=a or c or g or t
: NAME/KEY: misc_feature
: LOCATION: (60001)..(75000)
: OTHER INFORMATION: n=a or c or g or t
: NAME/KEY: misc_feature
: LOCATION: (75001)..(90000)
: OTHER INFORMATION: n=a or c or g or t
: NAME/KEY: misc_feature
: LOCATION: (90001)..(105000)
: OTHER INFORMATION: n=a or c or g or t
: NAME/KEY: misc_feature
: LOCATION: (105001)..(120000)
: OTHER INFORMATION: n=a or c or g or t

```

[illegible]

LOCATION: (480001)..(495000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (495001)..(510000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (510001)..(525000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (525001)..(540000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (540001)..(555000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (555001)..(570000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (570001)..(585000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (585001)..(600000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (600001)..(615000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (615001)..(630000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (630001)..(645000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (645001)..(660000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (660001)..(675000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (675001)..(690000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (690001)..(705000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (705001)..(720000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (720001)..(735000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (735001)..(750000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (750001)..(765000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (765001)..(780000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (780001)..(795000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (795001)..(810000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (810001)..(825000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (825001)..(840000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (840001)..(855000)

OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (855001)..(870000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (870001)..(885000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (885001)..(900000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (900001)..(915000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature

Query Match 3.4%; Score 55.8; DB 4; Length 1230025;
Best Local Similarity 49.9%; Pred. No. 0.00083;
Matches 197; Conservative 0; Mismatches 192; Indels 6; Gaps 2;

Qy 341 AGTACATTTTGAACATCCAGATGCTTGATATGATTTGATGACACAGGACTCTGACGA 400
Db 835626 ATTTAATTTGGATCATCCGAGCGCTTTGATATATGACTTATTAATTTGACATAAAC 835567
Qy 401 ACATCGTGAGGCGCAAAACGGTGAGGTCGACCTATGATTTGTGACA--CACTCAA 457
Db 83566 GTCATAAAATATGAGATTGTCGAAGCCAGTTTGTGATTTGTTAGTAATCGAT 835507
Qy 458 GGTTCACAGACACCGGTGTCTACCCGCGAGCGTGTCTGTTGAGGATCTTG 517
Db 835506 CTAAACGAGATGAGAAACGATTTATCCATCTAAAGTTATCTTGTTGAAGGATTTCTGG 835447
Qy 518 TGTCTTACAGCCAGAGATCCGGACATGTTCCACCTGCGCCTTCTGTGACACCGACT 577
Db 835446 TCTTTGAAATCAAGAACTTAGAGATCTTAGATATTTAGATCTTTGTAGAACCGATG 835387
Qy 578 CCGACGTACGCTGTCTCGAAGATTCTCCGGACGTG--GCCAGGAGGAGACCTGG 634
Db 835386 CTGATGAAGATATCAACGCCGTATGTTGAGATGTTCAAGAACAGAGATAGCGTGG 835327
Qy 635 AGCAGATTCGACGCGAGTACACCACTGTGTGAAGCGCGCTTCGAGGATTTGCGCTGC 694
Db 835326 ACTGCATCATGTCTCGTTATCTTTCTATGTTAAAGCCTATGCAATGAGAAATTTATGAGC 835267
Qy 695 CGACAAAGATATGCCGATGTGATCATCCACGA 729
Db 835266 CGACTCGAATATGCTGATATCATTTGTACATGGA 835232

RESULT 8
US-08-232-463-14/c
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEIFLINER, F.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:

```
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: pTZypc-Fls
US-08-232-463-14

Query Match
Best Local Similarity 3.3%; Score 54.4; DB 1; Length 7218;
Matches 10; Conservative 226; Mismatches 152; Indels 0; Gaps 0;

49 GAAAGGCGGGCGCGGACCCGATGCGCGAGCGGAGCGGAGTGGCTTCGCGCGGA 108
1434 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1375
109 GCGCAACACTGCGAGAGCCCGCGCGGAGCGGAGCGGAGTGGCTTCGCGCGG 168
1374 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1315
169 ATGAGGCGTGGCGCGGAGCGGAGCGGAGCGGAGTGGCTTCGCGGAGTATGAG 228
1314 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1255
229 TTGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 288
1254 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1195
289 AGGTTCAAGAGCTCTGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 348
1194 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1135
349 TTGACCATCCAGATGCTTGTGATGATGTTGATGACAGAGAGTGGAGAGATGCTG 408
1134 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1075
409 GAGGCGAAAAGCGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 436
1074 RRRRRRRATCGCAAGTCTCCCTGACCT 1047

RESULT 9
US-08-990-140-1
Sequence 1, Application US/08990140A
Patent No. 6093795
GENERAL INFORMATION:
APPLICANT: Olsen, Henrik S.
APPLICANT: Ruben, Steven M.
APPLICANT: Sonenberg, Nahum
APPLICANT: Method, Nathalie
APPLICANT: Rom, Eran
TITLE OF INVENTION: Human Prt1-like Subunit Protein (hprt1) and Human
FILE REFERENCE: 1488.070001
CURRENT APPLICATION NUMBER: US/08/990.140A
CURRENT FILING DATE: 1997-12-12
EARLIER APPLICATION NUMBER: US 60/033,151
```

```
EARLIER FILING DATE: 1996-12-13
NUMBER OF SEQ ID NOS: 13
SOFTWARE: Patentln Ver. 2.1
SEQ ID NO 1
LENGTH: 3032
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (97)..(2718)
US-08-990-140-1

Query Match
Best Local Similarity 3.3%; Score 53.6; DB 3; Length 3032;
Matches 86; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

4 GGGTCCTTCGACCTCGGCGCTGGGCGGCGCGCGCGCGCGCGCGCGCGCGCGCG 63
175 GAGCCGCGCGCAGCCGAGGGGCTGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCG 234
64 GGGACCCGATGCGCGGAGCGGAGCGCGGAGCGCGGATGCTTCGCGGAGCGGAG 123
235 GGGACCGAGGCTTCAGTGAAGAGTGGGAGTGGGAGCGGAGCGCGGAGCGGAG 294
124 AGCCCGCGCGCGGAGCGCGA 143
295 AGGACCGAGCGCGCGCGCGA 314

RESULT 10
US-09-546-238-1
Sequence 1, Application US/09546238
Patent No. 6316225
GENERAL INFORMATION:
APPLICANT: Olsen, Henrik S.
APPLICANT: Ruben, Steven M.
APPLICANT: Sonenberg, Nahum
APPLICANT: Method, Nathalie
APPLICANT: Rom, Eran
TITLE OF INVENTION: Human Prt1-like Subunit Protein (hprt1) Polynucleotides
FILE REFERENCE: 1488.070002
CURRENT APPLICATION NUMBER: US/09/546,238
CURRENT FILING DATE: 2000-04-10
PRIOR APPLICATION NUMBER: US 60/033,151
PRIOR FILING DATE: 1996-12-13
NUMBER OF SEQ ID NOS: 13
SOFTWARE: Patentln Ver. 2.1
SEQ ID NO 1
LENGTH: 3032
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (97)..(2718)
US-09-546-238-1

Query Match
Best Local Similarity 3.3%; Score 53.6; DB 4; Length 3032;
Matches 86; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

4 GGGTCCTTCGACCTCGGCGCTGGGCGGCGCGCGCGCGCGCGCGCGCGCGCGCG 63
175 GAGCCGCGCGCAGCCGAGGGGCTGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCG 234
64 GGGACCCGATGCGCGGAGCGGAGCGCGGAGCGCGGATGCTTCGCGGAGCGGAG 123
235 GGGACCGAGGCTTCAGTGAAGAGTGGGAGTGGGAGCGGAGCGCGGAGCGGAG 294
124 AGCCCGCGCGCGGAGCGCGA 143
295 AGGACCGAGCGCGCGCGCGA 314
```

```

QY      703   AAGTAGGCGCAATGTATCATCCACAGAGAGTGAACAATATGTTGCCATGCACCCGATC    762
           ||| |||| |
DB       146552 CAAATGGCGAATATTCTCATCTTGCTGTGTAATAAATCGCATTCATATATGTTA    146611
           ||| |||| |
QY      763   GTGCAGCACATCCAGACATCTCTGAAT    789
           ||| |||| |
DB       146612 AAAGCTCAAATCCTTCATTATTAAT    146638
           ||| |||| |

RESULT 12
US-09-643-990A-1
; Sequence 1, Application US/09643990A
; Patent No. 6528289
GENERAL INFORMATION:
APPLICANT: Robert D. Fleischmann
          Mark D. Adams
          Owen White
          Hamilton O. Smith
          J. Craig Venter
TITLE OF INVENTION: The Nucleotide sequence of
                    the Haemophilus influenzae Rd Genome, Fragments
                    Thereof, and Uses Thereof
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville,
STATE: MD
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: 3 1/2 inch diskette
COMPUTER: Dell Pentium
OPERATING SYSTEM: MS DOS v6.22
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/643,990A
FILING DATE: 23-Aug-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/487,429
FILING DATE: 1995-06-07
APPLICATION NUMBER: 08/426,787
FILING DATE: 1995-04-21
ATTORNEY/AGENT INFORMATION:
NAME: Kenley K. Hoover
REGISTRATION NUMBER: 40,302
REFERENCE/DOCKET NUMBER: PB186P1C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 301-610-5790
TELEFAX: 310-309-8439
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1830121 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-643-990A-1
Query Match              3.2%; Score 51.8; DB 4; Length 1830121;
Best Local Similarity    46.3%; Pred. No. 0.012;
Matches 207; Conservative 0; Mismatches 237; Indels 3; Gaps 1;

QY      346   AATTTCAGCATCCAGATGCGCTTGTAATATGATTGATGACACAGACTCTGAAGAATC    405
           ||| |||| |
DB       146192 AATTACGATCATCAAACTCTATGATGCGCATTTACTTATCCAACTTAATAAATCTA    146251
           ||| |||| |
QY      406   GTGAGGCGAAAACGGTGGAGGTGGCGGACCCTATGATTTTGTGACACACTCAAGGTTAC    465
           ||| |||| |
DB       146252 AAAATGGCAGTGCAGTAGTAGTGCTGCTTTATAGCTATGTAGAACATATCGCACCAAC    146311
           ||| |||| |

```


NAME: Kempier, Ph.D., Gail M
REGISTRATION NUMBER: 32,143
REFERENCE/DOCKET NUMBER: REG 070C
TELECOMMUNICATION INFORMATION:
TELEPHONE: 914-345-7400
TELEFAX: 914-345-7721
TELEX:
INFORMATION FOR SEQ ID NO: 106:
SEQUENCE CHARACTERISTICS:
LENGTH: 4092 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
FEATURE:
NAME/KEY: Coding Sequence
LOCATION: 200..3028
OTHER INFORMATION:
NAME/KEY: Human ROR2
LOCATION: 1...4092
OTHER INFORMATION:
US-08-469-537A-106

Query Match 3.1%; Score 50.4; DB 2; Length 4092;
Best Local Similarity 48.9%; Pred. No. 0.0015;
Matches 135; Conservative 0; Mismatches 141; Indels 0; Gaps 0;

QY 222 CATGAGTTGCTGGGACAGACGAGGTGAAACAGCGGAGGGAGGTGATCTGAG 281
DB 1552 CATGAAATGCCCTCATTAACGACACAAACAGCCAACTCAAAGATACGCTGTC 1611
QY 282 CCAGACAGGTTCTACAAAGTCTGACGCGACAGACGAGCCAAAGCCTTGAAGACA 341
DB 1612 TGCCGTGAGGTTCAATGAGGAGAGCTGGGAGAGACCGGTTGGAAAAGTCTAAGAGTCA 1671
QY 342 GTACAATTTTACCATCCAGATGCTTTGATAATGATTGTGACACAGGATCTGAAAGAA 401
DB 1672 CCGTTCGGCCCTGCGCCGCGGAGACAGCCAGGCTGTGCCATCAAAAGCTGAAGGA 1731
QY 402 CATGTTGAGGGCAAAACGTTGAGAGTCCGACCTATGATTTTGAACACATCAAGGTT 461
DB 1732 CAAAGCGAGGGGCCCCCTCGCGGAGAGTCCGCGATGAGCTATGCTCGACAGACGGCT 1791
QY 462 ACCAGAGACCAAGGTGTCTACCTGCGGAGGTGT 497
DB 1792 GCAACACCCCAACGTCGTCTGCTGGGCGTGT 1827

RESULT 15
US-07-642-734C-3
Sequence 3, Application US/07642734C
Patent No. 5824513
GENERAL INFORMATION:
APPLICANT: Katz, L
APPLICANT: Donadio, S
APPLICANT: Mcalpine, J B
TITLE OF INVENTION: Recombinant DNA Method for Producing
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: Edward H. Gorman
STREET: Abbott Laboratories D377/AP6D-2 One Abbott
CITY: Abbott Park
STATE: IL
COUNTRY: US
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/642,734C
FILING DATE: 17-JAN-91
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Danckers, Andreas M
REGISTRATION NUMBER: 32652
REFERENCE/DOCKET NUMBER: 4952.US.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 708-937-9396
TELEFAX: 708-938-2623
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 20235 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Saccharopolyspora erythraea
STRAIN: NRRL 238
FEATURE:
NAME/KEY: CDS
LOCATION: 19..10722
OTHER INFORMATION: /codon_start= 19
OTHER INFORMATION: /function= "gene eryA"
OTHER INFORMATION: /product= "eryA ORF2 encoding modules 3 & 4 for
OTHER INFORMATION: 6-deoxyerythronolide B"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 19..4470
OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: module 3"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 97..1482
OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: beta-ketoacyl ACP synthase of module 3"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1693..2670
OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: acyltransferase domain module 3"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 3406..3921
OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: beta-ketoreductase domain of module 3"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 4171..4428
OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: acyl carrier domain of module 3"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 4471..10722
OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: module 4"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 4471..5847
OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: beta-ketoacyl:Acrysinase domain of module"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 6054..7026
OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: acyltransferase domain of module 4"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 7165..9216
OTHER INFORMATION: /function= "approximate span of

```

1 OTHER INFORMATION: dehydratase and enoylreductase domains m
2 FEATURE:
3 NAME/KEY: misc_feature
4 LOCATION: 9433..9984
5 OTHER INFORMATION: /function= "approximate span
6 OTHER INFORMATION: beta-ketoreductase of module 4"
7 FEATURE:
8 NAME/KEY: misc_feature
9 LOCATION: 10225..10483
10 OTHER INFORMATION: /function= "approximate span of
11 OTHER INFORMATION: acyl carrier domain of module 4"
12 FEATURE:
13 NAME/KEY: CDS
14 LOCATION: 10723..20235
15 OTHER INFORMATION: /codon_start= 10723
16 OTHER INFORMATION: /function= "gene =eryA"
17 OTHER INFORMATION: /product= "orf3 encoding modules 5 & 6
18 OTHER INFORMATION: 6-deoxyerythronolide B formatio"
19 FEATURE:
20 NAME/KEY: misc_feature
21 LOCATION: 10723..15165
22 OTHER INFORMATION: /function= "approximate span of
23 OTHER INFORMATION: module 5"
24 FEATURE:
25 NAME/KEY: misc_feature
26 LOCATION: 10831..12174
27 OTHER INFORMATION: /function= "approximate span of
28 OTHER INFORMATION: beta-ketoacylACPSynthase domain of modul"
29 FEATURE:
30 NAME/KEY: misc_feature
31 LOCATION: 12379..13350
32 OTHER INFORMATION: /function= "approximatr span of
33 OTHER INFORMATION: acyltransferase domain of module 5"
34 FEATURE:
35 NAME/KEY: misc_feature
36 LOCATION: 14062..14610
37 OTHER INFORMATION: /function= "approximate span of
38 OTHER INFORMATION: beta-ketoreductase of module 5"
39 FEATURE:
40 NAME/KEY: misc_feature
41 LOCATION: 14857..15114
42 OTHER INFORMATION: /function= "approximate span of
43 OTHER INFORMATION: acyl carrier domain of module 5"
44 FEATURE:
45 NAME/KEY: misc_feature
46 LOCATION: 15166..20235
47 OTHER INFORMATION: /function= "approximate span of
48 OTHER INFORMATION: module 6"
49 FEATURE:
50 NAME/KEY: misc_feature
51 LOCATION: 15172..16569
52 OTHER INFORMATION: /function= "approximate span of
53 OTHER INFORMATION: beta-ketoacylACPSynthase domain of modul"
54 FEATURE:
55 NAME/KEY: misc_feature
56 LOCATION: 16768..17721
57 OTHER INFORMATION: /function= "approximate span of
58 OTHER INFORMATION: acyltransferase domain of module 6"
59 FEATURE:
60 NAME/KEY: misc_feature
61 LOCATION: 18379..18921
62 OTHER INFORMATION: /function= "approximate span of
63 OTHER INFORMATION: beta-ketoreductase domain of module 6"
64 FEATURE:
65 NAME/KEY: misc_feature
66 LOCATION: 19149..19398
67 OTHER INFORMATION: /function= "approximate span of
68 OTHER INFORMATION: acyl carrier domain of module 6"
69 FEATURE:
70 NAME/KEY: misc_feature
71 LOCATION: 19492..20235
72 OTHER INFORMATION: /function= "approximate span of
73 OTHER INFORMATION: thioesterase domain of module 6"

```

US-07-642-734C-3

Query Match	2.9%	Score 46.6;	DB 1;	Length 20235;
PostgreSQL Statistics	47.3%	Need No	0.034	

Best Local Similality: 47.28; Freq.NO. 0.034;
Matches 142; Conservative 0; Mismatches 159; Indels 0; Gaps 0;

33 GCGCGCCGGCCCGGGAGGGCGCGCGCGGACCCGATGCCGCGGAGCGGAGGCCGA 92

Db 5367 GGGCGTGC CGGCCG GACGTCGTGGAGGCGCACGGGCA CCGAGCTCGG 5426

93 GATGGCTTCGGCGGAGGCGAAGACTGCCGAGAGCCCCCGCCGGAGGCCGACCGTCCGCA 152

D5 5427 CGACCCGATCGAGGCCCGCCGCTCATCGGACCTACGGCCAGGACCGCGCCGCT 5486

153 CCAGCGCCCTCTCTGATAGGGTGAAGCGCGGCTGCCAGCGGAGTCGACCGTGTG 212

D8 548 / GCGGCLGGCLGGAGHAGCACACALGGCCACHCCTCAAGCCCCGGGCCGCGGC 559

[illegible]

22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 1041 1042 1043 1044 1045 1046 1047 1048 1049 1050 1051

[illegible]

0
1
2
3
4
5
6
7
8
9
A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
P
Q
R
S
T
U
V
W
X
Y
Z
[
\
]
^
_
`
a
b
c
d
e
f
g
h
i
j
k
l
m
n
o
p
q
r
s
t
u
v
w
x
y
z
{
|
}
~
_

—

Search completed: November 25, 2003, 02:03:52
Job time : 125.319 secs

Job time : 125.319 secs

Tr: - - - - -

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 24, 2003, 21:29:42 ; Search time 576.791 Seconds
(without alignments)
9210.067 Million cell updates/sec

Title: US-09-896-522-1

Perfect score: 1624
Sequence: 1 GTGGGGTGGCCCGAGCCGACCC.....CCAGGCTGTGGGGGCGAG 1624

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2172232 seqs, 163554964 residues

Total number of hits satisfying chosen parameters: 4344464

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA.*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1624	100.0	1624	9	US-09-896-522-1
2	1591	98.0	1648	10	US-09-833-381-2048
3	978.6	60.3	1288	14	US-10-037-270-546
4	880	54.2	9732	10	US-09-764-877-2804
5	880	54.2	19125	9	US-09-764-877-2805
6	834	51.4	834	9	US-09-896-522-3
7	485	29.9	510	12	US-10-023-386-2183
8	432.8	26.7	472	11	US-09-918-995-30379
9	377	23.2	1402	13	US-10-098-841-53
10	367.2	22.6	1310	10	US-09-925-300-220
11	321	13.6	472	11	US-10-066-543-1843
12	214.8	13.2	447	14	US-10-066-543-1843
13	210.4	13.0	455	11	US-09-918-985-23923
14	185	11.4	187	12	US-10-029-386-15883
15	175.6	10.8	2058	12	US-10-251-186-14
16	125	7.7	470	11	US-09-918-995-9667

17	106	6.5	1401	10	US-09-938-842A-1943	Sequence 1943, Ap
18	79.8	4.9	248	9	US-09-923-876-5506	Sequence 5906, Ap
19	68	4.2	222	10	US-09-764-877-399	Sequence 399, Ap
20	63	3.9	263	9	US-09-923-876-5130	Sequence 5130, Ap
21	61.6	3.8	3662	9	US-09-070-927A-215	Sequence 215, Ap
22	53.6	3.3	3032	9	US-09-954-043-1	Sequence 1, Appl
23	53.6	3.3	3032	14	US-10-200-154-1	Sequence 1, Appl
24	53.6	3.3	3152	10	US-09-880-107-3431	Sequence 3431, Ap
25	52.2	3.2	65	12	US-09-908-975-24159	Sequence 24159, Ap
26	51.8	3.2	1830121	14	US-10-329-960-1	Sequence 1, Appl
27	51.6	3.2	514	8	US-08-781-986A-955	Sequence 955, Ap
28	49.2	3.0	1392	14	US-10-156-761-4528	Sequence 4528, Ap
29	49.2	3.0	9025608	14	US-10-156-761-1	Sequence 1, Appl
30	47	2.9	926	9	US-09-770-445-403	Sequence 403, Ap
31	47	2.9	1188	10	US-09-938-842A-615	Sequence 615, Ap
32	46.4	2.9	338	9	US-09-770-791-869	Sequence 869, Ap
33	46.4	2.9	743	12	US-10-027-632-17673	Sequence 17673, A
34	46.4	2.9	743	13	US-10-027-632-17673	Sequence 17673, A
35	46.4	2.9	956	12	US-10-027-632-31508	Sequence 31508, A
36	46.4	2.9	956	13	US-10-027-632-31508	Sequence 31508, A
37	46	2.8	2561	10	US-09-976-740-48	Sequence 48, Appl
38	46	2.8	2561	13	US-10-023-529-48	Sequence 48, Appl
39	46	2.8	2561	13	US-10-023-523-48	Sequence 48, Appl
40	45.8	2.8	1449	14	US-10-156-761-3048	Sequence 3048, Ap
41	45.6	2.8	903	14	US-10-156-761-6755	Sequence 6755, Ap
42	45.6	2.8	9025608	14	US-10-156-761-1	Sequence 1, Appl
43	45.4	2.8	65	12	US-09-908-975-398	Sequence 398, Ap
44	45.4	2.8	5771	10	US-09-833-381-307	Sequence 307, Ap
45	45.2	2.8	253	10	US-09-974-300-626	Sequence 626, Ap

ALIGNMENTS

RESULT 1
US-09-896-522-1
; Sequence 1, Application US/09896522
; Parent No. US20020055161A1
; GENERAL INFORMATION:
; APPLICANT: Glucksmann, Maria A.
; TITLE OF INVENTION: 57658, A NOVEL HUMAN URIDINE KINASE AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: 38155200100
; CURRENT APPLICATION NUMBER: US/09/896,522
; CURRENT FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: 60/216,503
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Faastsq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1624
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (94)...(927)
US-09-896-522-1

Query Match	100.0%	Score 1624;	DB 9;	Length 1624;
Best local Similarity	100.0%	Pred. No. 0;		
Matches 1624;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;
QY	1	GTGGGGTGGCTTCGACCTCGGCGCTGGGCGGCGCGCGCGCGGGAAGGCGGCGGC	60	
DB	1	GTGGGGTGGCTTCGACCTCGGCGCTGGGCGGCGCGCGCGCGGGAAGGCGGCGGC	60	
QY	61	GGGGGGAACCCGATGGCGGGAAGCGAGGCGGATGGCTGGGCGGGAAGCGAAGCTGC	120	
DB	61	GGGGGGAACCCGATGGCGGGAAGCGAGGCGGATGGCTGGGCGGGAAGCGAAGCTGC	120	
QY	121	GAGAGCCCGCGCGCGAGGCGGACCGTCCGACCGAGCGGCGCTTCGATAGGGGTGAGC	180	
DB	121	GAGAGCCCGCGCGCGAGGCGGACCGTCCGACCGAGCGGCGCTTCGATAGGGGTGAGC	180	

QY 181 GCGGCGACCTGCCAGCGGGAAGTCGACCGCTGTGTGAGAAAGATCATGAGATTGCTGGACAG 240
 DB 181 GCGGCGACCTGCCAGCGGGAAGTCGACCGCTGTGTGAGAAAGATCATGAGATTGCTGGACAG 240
 QY 241 AACGAGGTGGAACAGCGGCGAGCGGAAGGTGTCATCTCGAGCCAGACAGATTCTACAG 300
 DB 241 AACGAGGTGGAACAGCGGCGAGCGGAAGGTGTCATCTCGAGCCAGACAGATTCTACAG 300
 QY 301 GTCCTGACGCGAGAGCAAGAGCGGAGGCTTGAAGAGACATCAATTTTGAACATCCA 360
 DB 301 GTCCTGACGCGAGAGCAAGAGCGGAGGCTTGAAGAGACATCAATTTTGAACATCCA 360
 QY 361 GATGCTCTTTGATATGATTTGATGACAGAGATCTGAAGAACATCTGTGGAGGGAACAG 420
 DB 361 GATGCTCTTTGATATGATTTGATGACAGAGATCTGAAGAACATCTGTGGAGGGAACAG 420
 QY 421 GTGAGAGTGCAGACCTATGATTTTGTGACACACTCAAGGTTTACAGAGACCAAGGTGTC 480
 DB 421 GTGAGAGTGCAGACCTATGATTTTGTGACACACTCAAGGTTTACAGAGACCAAGGTGTC 480
 QY 481 TACCTGCGGACGCTGTTCTGTGTTGAGGAGCATTTGGTGTCTACAGCCAGAGATCCGG 540
 DB 481 TACCTGCGGACGCTGTTCTGTGTTGAGGAGCATTTGGTGTCTACAGCCAGAGATCCGG 540
 QY 541 GACATGTTCCACCTGCGGCTCTTCTGTGACACCGACCTCCAGCTGAGGCTGTCTCAAGA 600
 DB 541 GACATGTTCCACCTGCGGCTCTTCTGTGACACCGACCTCCAGCTGAGGCTGTCTCAAGA 600
 QY 601 GTTCTCCGCGACGTCGCGCGGAGGAGGACCTGAGAGAGATTTGACGCGATACACCA 660
 DB 601 GTTCTCCGCGACGTCGCGCGGAGGAGGACCTGAGAGAGATTTGACGCGATACACCA 660
 QY 661 TTGCTGAAGCCGCGCTTTCGAGAGATTTGCTGCTGCGACAAAGAAATGCTCGATGTATC 720
 DB 661 TTGCTGAAGCCGCGCTTTCGAGAGATTTGCTGCTGCGACAAAGAAATGCTCGATGTATC 720
 QY 721 ATCCGAGAGAGAGGACAAATGTTGTCATGACCTGATCCGTCGACATCCAGAGAC 780
 DB 721 ATCCGAGAGAGAGGACAAATGTTGTCATGACCTGATCCGTCGACATCCAGAGAC 780
 QY 781 ATTCTGATGATGATCATCTGCAATGAGCAACGAGAGAGGTCATATGGGCGAGGTACAG 840
 DB 781 ATTCTGATGATGATCATCTGCAATGAGCAACGAGAGAGGTCATATGGGCGAGGTACAG 840
 QY 841 CGGACCTTTTCTGAGCCAGGAGACCACTGAGATGCTGACCTTGGCAAAACGCTCAT 900
 DB 841 CGGACCTTTTCTGAGCCAGGAGACCACTGAGATGCTGACCTTGGCAAAACGCTCAT 900
 QY 901 TTGAGATCCAGAGACCAACCCCACTGAGGAGGCTGCGAGCTCAAGGCAAGTCTCCGCC 960
 DB 901 TTGAGATCCAGAGACCAACCCCACTGAGGAGGCTGCGAGCTCAAGGCAAGTCTCCGCC 960
 QY 961 CGGATGATGATGATGAGGAGCTGAGGAGCGGAGCCCAACCACTGATGCTCTCTC 1020
 DB 961 CGGATGATGATGATGAGGAGCTGAGGAGCGGAGCCCAACCACTGATGCTCTCTC 1020
 QY 1021 GCGGACCCCGAGGAGGATGTTAGCAGGAGGCTTCTCACTCAGAGTGAATCTAGA 1080
 DB 1021 GCGGACCCCGAGGAGGATGTTAGCAGGAGGCTTCTCACTCAGAGTGAATCTAGA 1080
 QY 1081 TGTGTCCTGAGACTCAATTTGCTGAGACACTGACAGGCTTCTGAGGTTTTCAGCCAC 1140
 DB 1081 TGTGTCCTGAGACTCAATTTGCTGAGACACTGACAGGCTTCTGAGGTTTTCAGCCAC 1140
 QY 1141 TTAGGCTGCTGCTGCTTTAAAGATCCCTCTAGGTCATGAGAAATGCAAGAAATGTCGA 1200
 DB 1141 TTAGGCTGCTGCTGCTTTAAAGATCCCTCTAGGTCATGAGAAATGCAAGAAATGTCGA 1200
 QY 1201 GGAAGCTGAGAGGCTTCTGTGAGGAAATGAGGACATTAATTTGGGAAATTTGAGAGAC 1260
 DB 1201 GGAAGCTGAGAGGCTTCTGTGAGGAAATGAGGACATTAATTTGGGAAATTTGAGAGAC 1260

QY 1261 AGCTTAGACATGCGGCTGATGTTTGTGACAGTGAACCAACAGTGGAGAGATT 1320
 DB 1261 AGCTTAGACATGCGGCTGATGTTTGTGACAGTGAACCAACAGTGGAGAGATT 1320
 QY 1321 TTTTCCAGTGTGATCTGTTCTTACACACTCAACATTAATCAAAAGTTTGTGACA 1380
 DB 1321 TTTTCCAGTGTGATCTGTTCTTACACACTCAACATTAATCAAAAGTTTGTGACA 1380
 QY 1381 AGTACTTCTTTTACATGTTTACATGTCATGTCATGTTTCTGTTTCTGTTTCA 1440
 DB 1381 AGTACTTCTTTTACATGTTTACATGTCATGTCATGTTTCTGTTTCTGTTTCA 1440
 QY 1441 CAAGGCTGTTGAGCTCAAAACCTAATTTATGACCCAGTGTGAGTGCAGCGTG 1500
 DB 1441 CAAGGCTGTTGAGCTCAAAACCTAATTTATGACCCAGTGTGAGTGCAGCGTG 1500
 QY 1501 GCCTACACGAGATATGAGGAGCCACTGAGGAGATTTTCCCTTGTGCTTAAAG 1560
 DB 1501 GCCTACACGAGATATGAGGAGCCACTGAGGAGATTTTCCCTTGTGCTTAAAG 1560
 QY 1561 GCAGAGAGCGAGCGGATGCTTGAAGACCCAGATCAACCAAGCTTGTGCGGG 1620
 DB 1561 GCAGAGAGCGAGCGGATGCTTGAAGACCCAGATCAACCAAGCTTGTGCGGG 1620
 QY 1621 CCAG 1624
 DB 1621 CCAG 1624

RESULT 2
 US-09-833-381-2048
 ; Sequence 2048, Application US/09833381
 ; Patent No. US20020132090A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Robison, Keith E.
 ; TITLE OF INVENTION: No. US20020132090A1 Nucleic Acid and Protein Homologs
 ; FILE REFERENCE: 5800-119
 ; CURRENT APPLICATION NUMBER: US/09/833,381
 ; PRIOR FILING DATE: 2001-04-11
 ; PRIOR APPLICATION NUMBER: 09/516,448
 ; NUMBER OF SEQ ID NOS: 2050
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 2048
 ; LENGTH: 1648
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (1)..(1648)
 ; OTHER INFORMATION: n = A,T,C or G
 ; US-09-833-381-2048

Query Match 98.0%; Score 1591; DB 10; Length 1648;
 Best Local Similarity 99.6%; Pred. No. 0;
 Matches 1605; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

QY 14 CGACCTCGGCGTGGCGGCGCGGCGCGGCGGAGAGGCGCGGCGGAGCCCGAT 73
 DB 4 CGTCCCGCGCGCGCGCGGCGCGGCGCGGCGGAGAGGCGCGGCGGAGCCCGAT 63
 QY 74 GCGCGGAGACCGAGGCGGAGATGCTTGGCGGAGGCGCAAAATGCGAGACCCCGCGC 133
 DB 64 GCGCGGAGACCGAGGCGGAGATGCTTGGCGGAGGCGCAAAATGCGAGACCCCGCGC 123
 QY 134 CGGAGCGCAGCGTCCGACCAAGCGGCGCTTCTGTAAGGAGTGAAGCGGCGACTGCCA 193
 DB 124 CGGAGCGCAGCGTCCGACCAAGCGGCGCTTCTGTAAGGAGTGAAGCGGCGACTGCCA 183
 QY 194 GCGGGAAGTGCACCGTGTGAGAAAGATGAGAGTGTGCGGACGAAAGAGTGAAC 253
 DB 184 GCGGGAAGTGCACCGTGTGAGAAAGATGAGAGTGTGCGGACGAAAGAGTGAAC 243

Good Data

```

QY 254 AGCGGACGGGAAAGTGTGTCATCTGAGCCAGGACAGGTTCTCAAGAGTCTCGACGGCAG 313
Db 244 AGCGGACGGGAAAGTGTGTCATCTGAGCCAGGACAGGTTCTCAAGAGTCTCGACGGCAG 303
QY 314 AGCAGAAAGCCAAAGGCTTGAAAGAGACATATTTTGAACATTCAGATGCTTTGATA 373
Db 304 AGCAGAAAGCCAAAGGCTTGAAAGAGACATATTTTGAACATTCAGATGCTTTGATA 363
QY 374 ATGATTTGATGACAGAGATCTGTAAAGACATCTGTGAGGGCAAAACGCTGAGAGTCCGA 433
Db 364 ATGATTTGATGACAGAGATCTGTAAAGACATCTGTGAGGGCAAAACGCTGAGAGTCCGA 423
QY 434 CCTATGATTTTGTGACACACTCAAGTTACAGAGACAGAGTGTCTACCTGCGGACG 493
Db 424 CCTATGATTTTGTGACACACTCAAGTTACAGAGACAGAGTGTCTACCTGCGGACG 483
QY 494 TGGTTCGTGTTGAGGGATCTGTGTGTTCTACAGCCAGAGATCCGAGATGTTCCACC 553
Db 484 TGGTTCGTGTTGAGGGATCTGTGTGTTCTACAGCCAGAGATCCGAGATGTTCCACC 543
QY 554 TGGGCTCTTCTGTGACACCGACTCCGAGCTGCTCCAGAGTCTCCGGAGC 613
Db 544 TGGGCTCTTCTGTGACACCGACTCCGAGCTGCTCCAGAGTCTCCGGAGC 603
QY 614 TGGCCGAGGAGGAGCCTGAGAGATCTGACGAGATCAACACCTTGTGAGAGCGG 673
Db 604 TGGCCGAGGAGGAGCCTGAGAGATCTGACGAGATCAACACCTTGTGAGAGCGG 663
QY 674 CTTTCGAGAGTCTGCTGCGGCAAAAGAGTA-TGCCGATGTGATCATCCAGAGAG 732
Db 664 CTTTCGAGAGTCTGCTGCGGCAAAAGAGTATTCGATGTGATCATCCAGAGAG 723
QY 733 GTGAGCAATATGTGTGTCATCAACTGATGTGACAGACATCCAGACATTTGATGT 792
Db 724 GTGAGCAATATGTGTGTCATCAACTGATGTGACAGACATTTGATGT 783
QY 793 GACATCTGCAATATGAGCAGAGAGGCTCAATGGGGGAGCTACAAAGGAGACCTTTCT 852
Db 784 GACATCTGCAATATGAGCAGAGAGGCTCAATGGGGGAGCTACAAAGGAGACCTTTCT 843
QY 853 GAGCCAGGAGGAGCAGCCTGAGATGCTGACCTTGGCAAAAGCTCACATTTGAGTCCAGC 912
Db 844 GAGCCAGGAGGAGCAGCCTGAGATGCTGACCTTGGCAAAAGCTCACATTTGAGTCCAGC 903
QY 913 AGCAGACCCCACTGAGAGGCTGCCAGCCTCAAGGGAGGCTTCCGCGGACATGTGTGT 972
Db 904 AGCAGACCCCACTGAGAGGCTGCCAGCCTCAAGGGAGGCTTCCGCGGACATGTGTGT 963
QY 973 TCAGGAGCTGAGGCTGGGAGGCGCCACACCCACTGCTTCTCGGCGGACACCCAG 1032
Db 964 TCAGGAGCTGAGGCTGGGAGGCGCCACACCCACTGCTTCTCGGCGGACACCCAG 1023
QY 1033 GGGAGTGTGAGCAGAGGCTTCTCTACCTCAGAGAGTGAACCTAGATGTGCACTCAG 1092
Db 1024 GGGAGTGTGAGCAGAGGCTTCTCTACCTCAGAGAGTGAACCTAGATGTGCACTCAG 1083
QY 1093 ACTCAACTTCTGAGACACTGACAGGCTTCTGAGGTTTTCAGGCACTTAAAGCTGTG 1152
Db 1084 ACTCAACTTCTGAGACACTGACAGGCTTCTGAGGTTTTCAGGCACTTAAAGCTGTG 1143
QY 1153 CGGTTTAAAGATCCCTCTAGGTCACAGGAAATGCGCAGAAATGTGAGGAAAGCTGGGA 1212
Db 1144 CGGTTTAAAGATCCCTCTAGGTCACAGGAAATGCGCAGAAATGTGAGGAAAGCTGGGA 1203
QY 1213 GGGCTTGTGAGAGATGTGAGGACATTAATTTGGGAAATTTGAGAGAGCAGCTTACACT 1272
Db 1204 GGGCTTGTGAGAGATGTGAGGACATTAATTTGGGAAATTTGAGAGAGCAGCTTACACT 1263
QY 1273 GGGCTGCTGATTTTGTGAGACATGACACCAAGTGGGAGAGATTTTTCAGACTCTG 1332
Db 1264 GGGCTGCTGATTTTGTGAGACATGACACCAAGTGGGAGAGATTTTTCAGACTCTG 1323
QY 1333 ATCTGTTCTTACACTCAACATTACTCAAAAGTTTGTGGAACAAGTACTTCTT 1392

```

```

Db 1324 ATCTGTTCTTACACTCAACATTACTCAAAAGTTTGTGGAACAAGTACTTCTT 1383
QY 1393 TTTTACATGTTACATGTCCTCATGTTTCTGTTTCTGTTTCAACACAGGCTGTG 1452
Db 1384 TTTTACATGTTACATGTCCTCATGTTTCTGTTTCTGTTTCAACACAGGCTGTG 1443
QY 1453 TGGCCTTCAAAACCTTAATTTTCAATGACCCAGTGTGTTGATCCAGCGTGGCTTACAGGAT 1512
Db 1444 TGGCCTTCAAAACCTTAATTTTCAATGACCCAGTGTGTTGATCCAGCGTGGCTTACAGGAT 1503
QY 1513 ATGGGAGCCACTGAGAGATGTTTCCCTTGTGCTTGTGCTTAAAGCAGAGAGCA 1572
Db 1504 ATGGGAGCCACTGAGAGATGTTTCCCTTGTGCTTGTGCTTAAAGCAGAGAGCA 1563
QY 1573 GGGGATGCTTGTGAAGACCCAGACATCAACCCAGGCTTGTGGGGGCGAG 1624
Db 1564 GGGGATGCTTGTGAAGACCCAGACATCAACCCAGGCTTGTGGGGGCGAG 1615

RESULT 3
US-10-037-270-546
; Sequence 546, Application US/10037270
; Publication No. US20030104529A1
; GENERAL INFORMATION:
; APPLICANT: Yang, Y. Tom
; APPLICANT: Yang, Y. Tom
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Ping
; APPLICANT: Ma, Yunging
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhilwei
; APPLICANT: Tillinghast, John
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20030104529A1el Nucleic Acids and
; FILE REFERENCE: Polypeptides
; CURRENT APPLICATION NUMBER: US/10/037,270
; PRIOR FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: pc FL_genes Version 1.0
; SEQ ID NO 546
; LENGTH: 1288
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (95) .. (928)
US-10-037-270-546

Query Match 60.3%; Score 978.6; DB 14; Length 1288;
Best Local Similarity 99.6%; Pred. No. 3,4e-285;
Matches 981; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 29 GGGGGGCGCGGGGCGCGGGAAGGGGGGCGGGGAGCCGATGCGCGGAGGGAGG 88
Db 30 GGGTCCGCGCGGGCTGGGGAAGGGGGGCGGGGACCGGATGCGGGAGGGAGG 89
QY 89 CCGAGATGCTTCCGCGGAGGCGGAAGCTGCGAGAGCCCGCGCGGAGCGGACCGTC 148
Db 90 CCGAGATGCTTCCGCGGAGGCGGAAGCTGCGAGAGCCCGCGCGGAGCGGACCGTC 149

```

Good data

```

QY      149 CGCACACGCGGCGCTTCTCTGATAGGGGTGAGCGGCGGCACTGCCAGGGGAAAGTGCACCG 208
DB      150 CGCACACGCGGCGCTTCTCTGATAGGGGTGAGCGGCGGCACTGCCAGGGGAAAGTGCACCG 209
QY      209 TGTGTGAAAGATCATGAGTGTGCTGGGACAGAAAGAGTGGAAAGAGCGGACGGGAAAG 268
DB      210 TGTGTGAAAGATCATGAGTGTGCTGGGACAGAAAGAGTGGAAAGAGCGGACGGGAAAG 269
QY      269 TGTGTATCTTGGAGCCAGAGCAGGTTCTTCAAGAGTCTTGAACGGACAGAGCCAGAGG 328
DB      270 TGTGTATCTTGGAGCCAGAGCAGGTTCTTCAAGAGTCTTGAACGGACAGAGCCAGAGG 329
QY      329 CCTTGAAGAGCAGTATCAATTTTGAACATCCAGATGCTTGAATGAATTTGATGACACA 388
DB      330 CCTTGAAGAGCAGTATCAATTTTGAACATCCAGATGCTTGAATGAATTTGATGACACA 389
QY      389 GGAATCTGAAGAAATCGTGGAGGGCAAAAAGTGGAGGTGCGACCTATGATTTTGTGA 448
DB      390 GGAATCTGAAGAAATCGTGGAGGGCAAAAAGTGGAGGTGCGACCTATGATTTTGTGA 449
QY      449 CACACTCAAGGTTACAGAGACCAAGGTGTCTACCTGCGGACGTGGTCTGTTTGAAG 508
DB      450 CACACTCAAGGTTACAGAGACCAAGGTGTCTACCTGCGGACGTGGTCTGTTTGAAG 509
QY      509 GCATCTTGTGTGTCTTCAAGCCAGAGAGATCCGGGACATGTTCCACTGCGCTCTTCTGTG 568
DB      510 GCATCTTGTGTGTCTTCAAGCCAGAGAGATCCGGGACATGTTCCACTGCGCTCTTCTGTG 569
QY      569 ACACCGACTCCGAGCGTCAAGGTCTCTCGAAGAGTCTCCGGAGAGTCCGCGGAGGAGG 628
DB      570 ACACCGACTCCGAGCGTCAAGGTCTCTCGAAGAGTCTCCGGAGAGTCCGCGGAGGAGG 629
QY      629 ACTTGAAGCAGATTTCTGACGAGTACACCACTTCTGTGAAGCCGGCTTGTGAGAGTTCT 688
DB      630 ACTTGAAGCAGATTTCTGACGAGTACACCACTTCTGTGAAGCCGGCTTGTGAGAGTTCT 689
QY      689 GCGTCGCGGACAAAGAGATGCCGATGATCATCCACGAGAGTGGAGAAATTTGTTG 748
DB      690 GCGTCGCGGACAAAGAGATGCCGATGATCATCCACGAGAGTGGAGAAATTTGTTG 749
QY      749 CCATCAACCTGATGCTGAGCAGATCCAGAGATTTCTGAATGTGATCATCTGCAATGAGC 808
DB      750 CCATCAACCTGATGCTGAGCAGATCCAGAGATTTCTGAATGTGATCATCTGCAATGAGC 809
QY      809 ACCGAGAGAGGTCCAATGGGCGAGCTACAAAGCGAGCTTTTCTGAGCCAGGGAGCAACC 868
DB      810 ACCGAGAGAGGTCCAATGGGCGAGCTACAAAGCGAGCTTTTCTGAGCCAGGGAGCAACC 869
QY      869 CTGGGATGCTGACCTCTGGGCAAGCGGTCAATTTTGAAGTCCAGAGAGAGACCCCACTGAG 928
DB      870 CTGGGATGCTGACCTCTGGGCAAGCGGTCAATTTTGAAGTCCAGAGAGAGACCCCACTGAG 929
QY      929 GGGCTGCGAGCTCTCAGGGCAGGTCTCCCGCCGCGCATGTGTTCAGGAGCTGAGCCTG 988
DB      930 GGGCTGCGAGCTCTCAGGGCAGGTCTCCCGCCGCGCATGTGTGTTCAGGAACTGAGCCTG 989
QY      989 GGGAGCGCCACCCACACCCACTGCT 1013
DB      990 GGGAGCGCCACCCACACCCACTGCT 1014

```

```

RESULT 4
US-09-764-877-2804
; Sequence 2804, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005
; CURRENT APPLICATION NUMBER: US/09/764, 877
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper

```

```

; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2804
; LENGTH: 9732
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-877-2804

Query Match      54.2%; Score 880; DB 10; Length 9732;
Best Local Similarity 100.0%; Pred. No. 7.1e-255;
Matches 880; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      745 GTTGCCATCAACCTGATCGTGAGCAGCATCCAGACATTTCTGATGATGATCTGCAAA 804
DB      3901 GTTGCCATCAACCTGATCGTGAGCAGCATCCAGACATTTCTGATGATGATCTGCAAA 3960
QY      805 TGGCACCGAGAGGATGCAATGGGCGAGCTACAAAGCGAGCTTTTCTGAGCCAGGGAGC 864
DB      3961 TGGCACCGAGAGGATGCAATGGGCGAGCTACAAAGCGAGCTTTTCTGAGCCAGGGAGC 4020
QY      865 CACCTGGGATGCTGACCTCTGCGAAACGGTCAATTTGAGTCCAGCAGCAGACCCAC 924
DB      4021 CACCTGGGATGCTGACCTCTGCGAAACGGTCAATTTGAGTCCAGCAGCAGACCCAC 4080
QY      925 TGAAGGGGCTGCGAGGCTCAGGGCAGGTCTCCCGCCGAGATGTGTTCAGGACTGAG 984
DB      4081 TGAAGGGGCTGCGAGGCTCAGGGCAGGTCTCCCGCCGAGATGTGTTCAGGACTGAG 4140
QY      985 CTTGGGAGCGCCACCCACCAACCACTGCTTCTTGGGCGACCCAGGGGAGTGTAGC 1044
DB      4141 CTTGGGAGCGCCACCCACCAACCACTGCTTCTTGGGCGACCCAGGGGAGTGTAGC 4200
QY      1045 AGGAGAGCTTCTCTCACTCAGAGTGAAGAACTCAATGTGTCACTAAGTCACTAGTCT 1104
DB      4201 AGGAGAGCTTCTCTCACTCAGAGTGAAGAACTCAATGTGTCACTAAGTCACTAGTCT 4260
QY      1105 GGGACACTGACAGGCTTCTGAGGTTTCAAGCACTTGAAGCTGTGGGTTTAAAGAT 1164
DB      4261 GGGACACTGACAGGCTTCTGAGGTTTCAAGCACTTGAAGCTGTGGGTTTAAAGAT 4320
QY      1165 CCCTTAGGTCACCTGAAATGACAGAGATGACAGAGGCTGGAGGCTTCTGTGAG 1224
DB      4321 CCCTTAGGTCACCTGAAATGACAGAGATGACAGAGGCTGGAGGCTTCTGTGAG 4380
QY      1225 GAATGAGGCACTAATTTGGGAAATTGAGAGAGAGGCTTGAACCTGCGCTGAT 1284
DB      4381 GAATGAGGCACTAATTTGGGAAATTGAGAGAGAGGCTTGAACCTGCGCTGAT 4440
QY      1285 GTTTGTGACAGTGAACCAAGTGGAGAGGTTTTTTCAGTCTGATCTGTTCTTA 1344
DB      4441 GTTTGTGACAGTGAACCAAGTGGAGAGGTTTTTTCAGTCTGATCTGTTCTTA 4500
QY      1345 CACACTCAACACATTAATCAAAAGTTTGTGAACAAAGTACTTCTTTTAAATGTTA 1404
DB      4501 CACACTCAACACATTAATCAAAAGTTTGTGAACAAAGTACTTCTTTTAAATGTTA 4560
QY      1405 CATGTCTCATGTTTCTGTTTCTGTTTCAATTAACAAGGCTGTGTGGCTTCAAC 1464
DB      4561 CATGTCTCATGTTTCTGTTTCTGTTTCAATTAACAAGGCTGTGTGGCTTCAAC 4620
QY      1465 CTAAATTCATGACCCAGTGTGTTGAGTCCAGGAGGCTTACACGATATGAGGAGCCAC 1524
DB      4621 CTAAATTCATGACCCAGTGTGTTGAGTCCAGGAGGCTTACACGATATGAGGAGCCAC 4680
QY      1525 TGAAGGATGTTTCCCGCTTGTGTGCTTAAAGCAGAGAGAGGAGCGAGTCCCT 1584
DB      4681 TGAAGGATGTTTCCCGCTTGTGTGCTTAAAGCAGAGAGAGGAGCGAGTCCCT 4740
QY      1585 GGAAGCACCCAGATACACCCAGGCTTGTGGGGGCCAG 1624
DB      4741 GGAAGCACCCAGATACACCCAGGCTTGTGGGGGCCAG 4780

```

RESULT 5
US-09-764-877-2805
; Sequence 2805, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005
; CURRENT APPLICATION NUMBER: US/09/764,877
; PRIOR FILING DATE: 2001-01-17
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2805
; LENGTH: 19125
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-877-2805

Query Match 54.2%; Score 880; DB 10; Length 19125;
Best Local Similarity 100.0%; Pred. No. 1e-254;
Matches 880; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 745 GTTGCCATCAACCTGATCGTGCAGACATCCAGACATTCGAATGATCTGCAAA 804
Db 5997 GTTGCCATCAACCTGATCGTGCAGACATCCAGACATTCGAATGATCTGCAAA 6056
Qy 805 TGGCACCAGAGAGGCTCAATGGCGGAGCTCAAGCGGACCTTTTCTGACCGAGGAG 864
Db 6057 TGGCACCAGAGAGGCTCAATGGCGGAGCTCAAGCGGACCTTTTCTGACCGAGGAG 6116
Qy 865 CACCCCTGGAGTGTGACCTCTGCGAAACGTCAATTTGAGTCCAGACAGACCCAC 924
Db 6117 CACCCCTGGAGTGTGACCTCTGCGAAACGTCAATTTGAGTCCAGACAGACCCAC 6176
Qy 925 TGAAGGCTCCGAGCTCTGAGGAGGCTCCGCGCGGATGTTTGAAGGAGCTGAG 984
Db 6177 TGAAGGCTCCGAGCTCTGAGGAGGCTCCGCGCGGATGTTTGAAGGAGCTGAG 6236
Qy 985 CTTGGGAGCCGCCCAACCAACCACTGCTCTCTCGGCGACCCAGGAGGAGTTTAC 1044
Db 6237 CTTGGGAGCCGCCCAACCAACCACTGCTCTCTCGGCGACCCAGGAGGAGTTTAC 6296
Qy 1045 AGGAGGCTCTCTCACTGAGAGTGAATCTGAGTGTCACTGAGACTCAACTTGTCT 1104
Db 6297 AGGAGGCTCTCTCACTGAGAGTGAATCTGAGTGTCACTGAGACTCAACTTGTCT 6356
Qy 1105 GGGACACTGACAGGCGTTCTGAGAGTTTTCAGCCACTTAAAGCTGCTGCTTAAAGAT 1164
Db 6357 GGGACACTGACAGGCGTTCTGAGAGTTTTCAGCCACTTAAAGCTGCTGCTTAAAGAT 6416
Qy 1165 CCCTCTAGTCACTGAGAAATGCCACAGAAATGTGACAGAAAGCTGGAGGCTTCTGTAG 1224
Db 6417 CCCTCTAGTCACTGAGAAATGCCACAGAAATGTGACAGAAAGCTGGAGGCTTCTGTAG 6476
Qy 1225 GAATGTGAGGACATTTTGGGGAATTTGAGAGACAGCCTTAACACTGCTGCTGAT 1284
Db 6477 GAATGTGAGGACATTTTGGGGAATTTGAGAGACAGCCTTAACACTGCTGCTGAT 6536
Qy 1285 GTTTTGTGACAGTAAACCAAGTGGAGAGAGTTTTCAGTCTGATCTGATTTTAA 1344
Db 6537 GTTTTGTGACAGTAAACCAAGTGGAGAGAGTTTTCAGTCTGATCTGATTTTAA 6596
Qy 1345 CACACTCACACATAAATCAAAAAGTTTGTGAACAAGTACTTCTTTTACATGTA 1404
Db 6597 CACACTCACACATAAATCAAAAAGTTTGTGAACAAGTACTTCTTTTACATGTA 6656
Qy 1405 CAGTCTCATGTTTCTGTTTCTGTTTCAATACAAAGCTGTTGTTGCTTAAAC 1464
Db 6657 CAGTCTCATGTTTCTGTTTCTGTTTCAATACAAAGCTGTTGTTGCTTAAAC 6716
Qy 1465 CTAATTTTCATGACCAAGTGTGTCAGTCCAGGCTGCTTACAGATATGAGGAGCCAC 1524
|||||

Db 6717 CTAATTTTCATGACCAAGTGTGTCAGTCCAGGCTGCTTACAGATATGAGGAGCCAC 6776
Qy 1525 TGAAGATGTTTCCCTTCTGTTTCTTAAAGCAGAAAGGAGCGGATCCCT 1584
Db 6777 TGAAGATGTTTCCCTTCTGTTTCTTAAAGCAGAAAGGAGCGGATCCCT 6836
Qy 1585 GGAAGCACCAGACATACACCCAGGCTTGGCGGCGCAG 1624
Db 6837 GGAAGCACCAGACATACACCCAGGCTTGGCGGCGCAG 6876

RESULT 6
US-09-896-522-3
; Sequence 3, Application US/09896522
; Patent No. US20020055161A1
; GENERAL INFORMATION:
; APPLICANT: Glucksmann, Maria A.
; TITLE OF INVENTION: 57658, A NOVEL HUMAN URIDINE KINASE AND
; FILE REFERENCE: 381552001700
; CURRENT APPLICATION NUMBER: US/09/896,522
; PRIOR FILING DATE: 2001-06-28
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 834
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-896-522-3

Query Match 51.4%; Score 834; DB 9; Length 834;
Best Local Similarity 100.0%; Pred. No. 1.5e-241;
Matches 834; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 94 ATGCTTCGCGGAGGAGGAGAGCTGCGAGAGCCCGCGCGGAGCGGACCGTCCGAC 153
Db 1 ATGCTTCGCGGAGGAGGAGAGCTGCGAGAGCCCGCGCGGAGCGGACCGTCCGAC 60
Qy 154 CAGCGGCCCTTCTCTGTAAGAGGCTGAGCGGCGGCTCCAGCGGAGAGTGCAGCTGTGT 213
Db 61 CAGCGGCCCTTCTCTGTAAGAGGCTGAGCGGCGGCTCCAGCGGAGAGTGCAGCTGTGT 120
Qy 214 GAGAATATATGAGGCTGCTGAGCAGAAAGAGTGAACAGCGGAGCGGAGAGTGTCT 273
Db 121 GAGAATATATGAGGCTGCTGAGCAGAAAGAGTGAACAGCGGAGCGGAGAGTGTCT 180
Qy 274 ATCTGAGCAGAGAGGCTTCTAAGAGTCTGAGCGGAGAGAGGCAAGGCTTG 333
Db 181 ATCTGAGCAGAGAGGCTTCTAAGAGTCTGAGCGGAGAGAGGCAAGGCTTG 240
Qy 334 AAAGCAGTACATTTTGAACATCCAGATGCTTGTATATGATTTGATGACAGAGCT 393
Db 241 AAAGCAGTACATTTTGAACATCCAGATGCTTGTATATGATTTGATGACAGAGCT 300
Qy 121 GAGAAATATGAGGCTGCTGAGCAGAAAGGAGTGAACAGCGGAGCGGAGAGTGTCT 180
Qy 274 ATCTGAGCAGAGAGGCTTCTAAGAGTCTGAGCGGAGAGAGGCAAGGCTTG 333
Db 181 ATCTGAGCAGAGAGGCTTCTAAGAGTCTGAGCGGAGAGAGGCAAGGCTTG 240
Qy 334 AAAGCAGTACATTTTGAACATCCAGATGCTTGTATATGATTTGATGACAGAGCT 393
Db 241 AAAGCAGTACATTTTGAACATCCAGATGCTTGTATATGATTTGATGACAGAGCT 300
Qy 394 CTGAGAACATGCTGAGGAGGCAAAAGCTGAGAGTGCAGCCTATGATTTTGTGACAC 453
Db 301 CTGAGAACATGCTGAGGAGGCAAAAGCTGAGAGTGCAGCCTATGATTTTGTGACAC 360
Qy 454 TCAAGGTTTCAAGAGCAGCGGTGCTACCTGCGGAGAGTGTCTGTTTGAAGGCAATC 513
Db 361 TCAAGGTTTCAAGAGCAGCGGTGCTACCTGCGGAGAGTGTCTGTTTGAAGGCAATC 420
Qy 514 TTGTTTCTTACAGCAGGAGATCCGAGCATGTTTCACTGCGCTCTTCTGTGACAC 573
Db 421 TTGTTTCTTACAGCAGGAGATCCGAGCATGTTTCACTGCGCTCTTCTGTGACAC 480
Qy 574 GACTCCGAGCTAGGCTGTCTGAGAGATTTCTCGGAGAGTGTGCGGAGAGGAGCTG 633
Db 481 GACTCCGAGCTAGGCTGTCTGAGAGATTTCTCGGAGAGTGTGCGGAGAGGAGCTG 540
Qy 634 GAGCAGATTTCTGACGAGTACACACTTCTGTGAAGCGGCTTGTGAGGAGTTCTGCTG 693
|||||

```
Db 541 GAGCAGATTGAGCGACGACCACTTCTGTAAGCCGGCTTGAGAGTTGCTGCTG 600
Qy 694 CCGAACAAGAGTATCCCGATGTGTATATCCACAGAGATGAGACAAATATGTTGCCATC 753
Db 601 CCGAACAAGAGTATGCGATGTGTATATCCACAGAGATGAGACAAATATGTTGCCATC 660
Qy 754 AACCTGATGTCAGACATCCAGACATTTCTGAATGTTGACATCTGCAATATGACACCA 813
Db 661 AACCTGATGTCAGACATCCAGACATTTCTGAATGTTGACATCTGCAATATGACACCA 720
Qy 814 GAGAGTCCAAATGAGCGAGCTACAAAGCGACCTTTTCTGAGCCAGAGGACCACTGGG 873
Db 721 GAGAGTCCAAATGAGCGAGCTACAAAGCGACCTTTTCTGAGCCAGAGGACCACTGGG 780
Qy 874 ATGCTGACCTCTGCAACCGTCACTTTGAGATGTCAGACACCACTGAG 927
Db 781 ATGCTGACCTCTGCAACCGTCACTTTGAGATGTCAGACACCACTGAG 834
```

RESULT 7

```
US-10-029-386-2183
; Sequence 2183, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AECOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 2183
; LENGTH: 510
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR9,1
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.5
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.8
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.6
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.7
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.5
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.7
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.3
; OTHER INFORMATION: NT HIT: g11899252, EVALUATE 0.00e+00
; OTHER INFORMATION: EST HUMAN HIT: BF664526.1, EVALUATE 0.00e+00
; OTHER INFORMATION: SWISSPROT HIT: P52623, EVALUATE 2.00e-25
US-10-029-386-2183
```

```
Query Match 29.9%; Score 485; DB 12; Length 510;
Best Local Similarity 100.0%; Pred. No. 5.1e-136;
Matches 485; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 745 GTTGCCATCAACCTGATGTGACAGACATCCAGACATTTCTGAATGTTGACATTCGAAA 804
Db 26 GTTGCCATCAACCTGATGTGACAGACATCCAGACATTTCTGAATGTTGACATTCGAAA 85
Qy 805 TGGACCGAGAGAGGCTCAATGAGGAGGAGCTACAAAGCGACCTTTTCTGAGCCAGAGGAG 864
Db 86 TGGACCGAGAGAGGCTCAATGAGGAGGAGCTACAAAGCGACCTTTTCTGAGCCAGAGGAG 145
Qy 865 CACCTGAGATGCTGACCTTGGCAACGATCAATTTGAGTCCAGACAGACCCAC 924
Db 146 CACCTGAGATGCTGACCTTGGCAACGATCAATTTGAGTCCAGACAGACCCAC 205
Qy 925 TGAAGGCTGCGGAGCTCAAGGCAAGTCTCCCGCCGCGCATGTGTCTTCAAGGACCTGAG 984
Db 206 TGAAGGCTGCGGAGCTCAAGGCAAGTCTCCCGCCGCGCATGTGTCTTCAAGGACCTGAG 265
```

```
Qy 985 CTTGGGAGGCCCAACCCACACCACTTCTCTGCGGCGCACCCAGGGAGGTTAGC 1044
Db 266 CTTGGGAGGCCCAACCCACACCACTTCTCTGCGGCGCACCCAGGGAGGTTAGC 325
Qy 1045 AGCGAGGCTTCTCTCACTCAAGAGTGAACCTGATGTGTCACTGACCTCAACTTCTCT 1104
Db 326 AGCGAGGCTTCTCTCACTCAAGAGTGAACCTGATGTGTCACTGACCTCAACTTCTCT 385
Qy 1105 GGGACACTGACAGGCTTCTCTGAGTGTTCAGCCACTTGAAGCTTTCGGTTTAAAGAT 1164
Db 386 GGGACACTGACAGGCTTCTCTGAGTGTTCAGCCACTTGAAGCTTTCGGTTTAAAGAT 445
Qy 1165 CCGCTGATGCTGAGAAATGCAAGATGTGACAGAAAGCTTGGAGAGCTTCTGTAG 1224
Db 446 CCGCTGATGCTGAGAAATGCAAGATGTGACAGAAAGCTTGGAGAGCTTCTGTAG 505
Qy 1225 GAATG 1229
Db 506 GAATG 510
```

RESULT 8

```
US-09-918-995-30379
; Sequence 30379, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 30379
; LENGTH: 472
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(472)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-30379
```

```
Query Match 26.7%; Score 432.8; DB 11; Length 472;
Best Local Similarity 98.2%; Pred. No. 3.1e-120;
Matches 437; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
```

```
Qy 177 GAGCGCGGACACTGCCAGCGGGAAGTCGACCCGTGTGAGAAATATATGAGTTGCTGG 236
Db 28 GTGCGTGGAAAGCGCAGCGGGAAGTCGACCCGTGTGAGAAATATATGAGTTGCTGG 87
Qy 237 ACAGAACGAGAGTGAACAGCGGAGCGGAAGGTGATCTCTGAGCAGACAGTTCTA 296
Db 88 ACAGAACGAGAGTGAACAGCGGAGCGGAAGGTGATCTCTGAGCAGACAGTTCTA 147
Qy 297 CAAGGTCTTGACGCGCAGAGCAGAAAGCCCTTGAAGACAGTACAAATTTTGACA 356
Db 148 CAAGGTCTTGACGCGCAGAGCAGAAAGCCCTTGAAGACAGTACAAATTTTGACA 207
Qy 357 TCCAGATGCTTGTGATTAATGATTTGACACAGAGATCTTGAAAGACATGTGAGGGCAA 416
Db 208 TCCAGATGCTTGTGATTAATGATTTGACACAGAGATCTTGAAAGACATGTGAGGGCAA 267
Qy 417 AACGGTGAAGGTCCGACCTATGATTTTGTGACACACTAAAGTTACAGAGACACGGT 476
Db 268 AACGGTGAAGGTCCGACCTATGATTTTGTGACACACTAAAGTTACAGAGACACGGT 327
Qy 477 GGTCTACCTGCGAGAGTGGTTCTGTTTGAAGGCACTTGTGTTTACAGCAGAGAT 536
```

Db 328 GGTCTACCTCGAGACGTGTTCTGTTTGAAGGCATCTGTGTTTCTACAGCCAGAGAT 387
Qy 537 CCGGACATTTTCCACTCGTGGCCCTTTCGTGGACACCGACTCCGACGTGAGGTGTTCG 536
Db 388 CCGGACATTTTCCACTCGTGGCCCTTTCGTGGACACCGACTCCGACGTGAGGTGTTCG 447
Qy 597 AAGATTCTCCGGACGTGGCCGA 621
Db 448 AAGATTCTCCGGACGTGGCCGA 472

RESULT 9
US-10-098-841-53
; Sequence 53, Application US/10098841
; Publication No. US20020197679A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Xu, Chongjun
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yungqing
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhao, Qing A.
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: Wehrman, Tom
; APPLICANT: Zhang, Jie
; APPLICANT: Qian, Xiaohong B.
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20020197679A1el Nucleic Acids and
; FILE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2
; CURRENT APPLICATION NUMBER: US/10/098, 841
; PRIOR FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: 09/598, 042
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/552, 317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488, 725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 331
; SOFTWARE: pc_fl_genes Version 1.0
; SEQ ID NO 53
; LENGTH: 1402
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (290)..(1075)
US-10-098-841-53

Query Match 23.2%; Score 377; DB 13; Length 1402;
Best Local Similarity 69.0%; Pred. No. 4.4e-103; Indels 5; Gaps 2;
Matches 546; Conservative 0; Mismatches 240;

Qy 22 GCGCTGGGCGGCGCGCGCGGAGAGGAGGCGCGCGCGGACCCGATCGCGGGA 81
Db 211 GCGGAGGAGATCCGACCGGCGCGGCGGAGCGTGGCTTCGTCACAGGACGCG 270
Qy 82 GCGGAGGCGGAGATGCTTGGCGGAGCGGAGACTCGGAGAGCCCGCGGAGGCC 141
Db 271 GAGGAGGCGGCGGCGGACCAATGCGCGGAGCAGCGACAGACCCCTGAGAACCAAGCA 330
Qy 142 GACCGTCCGACAGCGGCGCTTCTGATAGGGGTGAGCGGCGGCACTGCGACAGGAG 201
Db 331 GCCCAACGCGGCG--GAGCCCTTCTTATAGCGCTCAGCCGGGAAACAGTTAGCGGAG 388
Qy 202 TCACCGGTGTGAGAGATCATGAGTTGCTGGGACAGAACGAGGTGAAACGCGGAG 261
Db 389 TCTTCGTGTGTCTAAGATCGTCAAGCTCTGCGGCGAGAAATGAGGTGAGCTATCGCGAG 448

Qy 262 CGAAGGTGATCTCTAGCCAGAGCAGGTTCTACAAGTCTCTGACGCGAGACGAAG 321
Db 449 AAGCAGGTGTATCTCTAGCCAGAGATGACCTTCTACGTTCTTACCTCGAGGAGAA 508
Qy 322 GCCAAGGCTTGAAGAGACATCAATTTTGAACCATCCAGATGCTTTGATATGATTG 381
Db 509 GCCAAGGCTTGAAGGCGGACATTTTGAACCATCCGATGCTTTGATATGAACTC 568
Qy 382 ATGCACAGACCTTGAAGAACATGTCGAGGCGAAACGCTGAGGTGCCGACTATGAT 441
Db 569 ATTCTCAAAACACTCAAGAAATCACTGAAGGGAAGAAACAGTCAGATCCCGGTATGAC 628
Qy 442 TTTGTACACACTCAAGTTTACAGAGACACGAGTGTCTTACCTTCGAGAGGTTTCG 501
Db 629 TTTGTCTCCATTCCTCGAAGGAGAGAGATTACTGTCTATCCGCAACCGTGTCTC 668
Qy 502 TTTGAGGACCTTGTGTTCTTACAGCCAGAGATCCGGAACATGTTCCAGCTGCGCTC 561
Db 689 TTTGAGGATCTGTGCTTCTTACTCTCCAGAGGTACGAGACTGTTCCAGATGAAGCTT 748
Qy 562 TTCTGAGACACGACTCCGACGTCAGGCTGTCTCGAAGAGTTCTCGGACGTCGC-- 618
Db 749 TTTGTGATACAGATGCGGACACCGGCTCTCAGCGAGATTTAAGGACATCAGCGAG 808
Qy 619 CGAGGAGGAGCCTGAGAGCATTTGACGACATACACACTTCTGTGAAGCGGCTTC 678
Db 809 AGAGGCGAGATCTTGAAGCATTTTATCTCATTAAGTTCGTAAGCTTGTCTT 868
Qy 679 GAGAGTTCTGCTGCGCGCAAGAAAGTATGCGATGTATCATCCACGAGAGTGAG 738
Db 869 GAGGATTTCTGCTGCGCAAGAAAGTATGCTGATGTATCATCTCTAGAGGTGAGAT 928
Qy 739 AATATGTTCCATCAACTGATGTCGAGCAGCATTCAGAGCATTTGATGATGATC 798
Db 929 AATCTGTGCGCATCAACTGATGTCGAGCAGCATTCAGAGCATTTGATGATGATG 988
Qy 799 TCGAATGCA 809
Db 989 TCGAATGCA 999

RESULT 10
US-09-925-300-220
; Sequence 220, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben,
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 220
; LENGTH: 1310
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-300-220

Query Match 22.6%; Score 367.2; DB 10; Length 1310;
Best Local Similarity 68.6%; Pred. No. 3.9e-100; Indels 7; Gaps 3;
Matches 550; Conservative 1; Mismatches 244;

Qy 14 CGAAGTGGGCTGGGCGGCGCGCGGCGGAGGAGGCGGCGGAGACCCGAT 73
Db 140 CGGCTGGGAAAGCGAGGAGATCCGACGCGGCGCGGCGGAGCGTGTCTCGGTTGG 199

```

QY 74 GCGCGGAGCGGAGCGGAGTGGCTTCGGCGGAGCGGAGTGCAGAGACCCCGGC 133
    |||
Db 200 CACAGGACCGGAGGAGGAGGCGCGGCAACATGGCCGGGAGACGACGAGACCCCTGC 259
    |||
QY 134 CGAGGCGGACCGCTCCGACCAAGC--GCCCTTCTGATAGGGGTGAGCGCGGACTG 190
    |||
Db 260 AGAACCACACAGCGCCCAACCGCGGAGCCCTTCTTATAGGCTCA-CGGGGAGACAG 318
    |||
QY 191 CCAGCGGAGAGTGCACCGGTGTGTGAGAGATCATGGATGTGGGACAGAACGAGGTGG 250
    |||
Db 319 CTAGCGGAGAGTCTTCCGTGTGTCTAGATGTGTACACTCTGGGAGAGATGAGGTGG 378
    |||
QY 251 AACGCGGACGAGAGGTGTCTATCTGAGCGAGAGAGGTCTTCAAGGTCTCTGACGG 310
    |||
Db 379 ACTATCGCAGAGAGGTGTCTATCTGAGCGAGAGATGAGTCTTCAAGGTCTCTTACT 438
    |||
QY 311 CAGAGCAGAGGCGCAGGCTTGAAGAGACAGTACATTTTGAACCATCCATGCTCTTGG 370
    |||
Db 439 CGGAGCAGAGGCGCAGGCTTGAAGAGGSCAGTTCACCTTTGACCAACCGGATGCTTTG 498
    |||
QY 371 ATTAATGATTTGATGACAGAGCTCTGAGAAATCTGTGAGGGGCAAAACGGTGGAGGTGC 430
    |||
Db 499 ACAATGASTCTTTTCAAAACCTCAAGAAATCATGAGGAGAAACAGTCCAGATCC 558
    |||
QY 431 CGACCTATGATTTGTGACACACTCAAGTTACAGAGACCAAGGTGTCTTACCTCGCG 490
    |||
Db 559 CCGGTATGACTTTGTCTCCATTTCCGGAAAGAGAGACAGTTACTGTCTATCCCGAG 618
    |||
QY 491 ACGTGTCTTGTGAGGCACTTTGTGTCTTACAGCGAGAGATCCGGGACATGTTCC 550
    |||
Db 619 ACGGTGTCTTGTGAGGCACTTTGTGTCTTACAGCGAGAGATCCGGGACATGTTCC 678
    |||
QY 551 ACCTGCGGCTCTGTGAGACCCGACCTCGAGGCTGTCTGAGAGGTCTTCCGGG 610
    |||
Db 679 AGATGAGACTTTTGTGATACAGATGCGGACACCGGCTCTCACGAGATTTAAAGG 738
    |||
QY 611 ACGTCCGC--GAGAGGAGGAGCTGAGAGATTTGAGCAGTACCAACCTTCTGTA 667
    |||
Db 739 ACATCAGGAGAGAGGAGGAGTCTTGAGCAGATTTTATCTCAGTACATTCGTTGCA 798
    |||
QY 668 AGCCGCGCTTGGAGAGTTCGCTGCGGACAAAGAAATAGCCGATGTATCTCCAC 727
    |||
Db 799 AGCTGCTTTGAGAAATTCGCTTGCCCAAAAGAAATAGCTGATGTATCTCCCTA 858
    |||
QY 728 GAGAGTGCACATATGTTGCTCATCACTGATGTGTGAGGATCATCCAGATTTCTGA 787
    |||
Db 859 GAGGTGCAGATATCTGTGCTGCTCATCACTCATGTGTGACACATCCAGATCTCTGA 918
    |||
QY 788 ATGTGACATCTGCAATGCA 809
    |||
Db 919 ATGAGGGGCTCCAAACGCA 940
    |||

RESULT 11
US-09-918-995-24042
; Sequence 24042, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 24042
; LENGTH: 472
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:

```

23

```

; NAME/KEY: misc.feature
; LOCATION: (1) -- (472)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-24042

Query Match      13.6%; Score 221; DB 11; Length 472;
Best Local Similarity 73.0%; Pred. No. 4e-56;
Matches 284; Conservative 0; Mismatches 105; Indels 0; Gaps 0;

QY 159 GCCCTTCTGATAGGGGTGAGCGCGGACATGCGCAGCGGGAAGTCCAGCGTGTGAGAA 218
    |||
Db 84 GCCCTTCTTATAGGCGTCAAGCGGGGAAACAGTACGGGCAAGTCTTCCGTGTGTCTAA 143
    |||
QY 219 GATCATGAGTTCCTGGGACAGAACGAGTGAACAGCGGACAGGAGGTGATCTCT 278
    |||
Db 144 GATGTGACACTCTCTGGGAGAGATGAGTGTGACTATCGCAGAACAGAGGTGATCTCT 203
    |||
QY 279 GAGCAGAGACAGGTCTTCAAGGTCTTGAAGGAGAGCAGAACGAGCCCTTGAAGG 338
    |||
Db 204 GAGCAGAGATAGCTTCTACCGGTCTTACCTCGAGAGAGAGCCAAAGCCCTGAAGGG 263
    |||
QY 339 ACAGTACATTTTGAACCATCAGATGCTTTGATTAATGATTTGATGACAGACTCTGAA 358
    |||
Db 264 CAGTTCACCTTTGACACCCGATGCTTTGACAAATGAACTCATTTCTCAAAACACTCAA 323
    |||
QY 399 GAACATGTGAGAGGCAAAACGGTGAAGTGGCCGACCTATGATTTGTGACACACTCAAG 458
    |||
Db 324 AGAATCTAGAGAGGAAACAGTCCAGATCCCGTGTATGATCTTGTCTCCATTTCCG 383
    |||
QY 459 GTTACAGAGACCAAGGTGTCTACCTTGGAGCGTGTCTTGTGAGGCAATCTTGGT 518
    |||
Db 384 GAAGAGAGACAGATGATCTATATCCCGCAGAGTGTGTCTTTGAAGGATCTCTGGC 443
    |||
QY 519 GTTCTACAGCCAGAGATCCGGGACATGT 547
    |||
Db 444 CTCTACTCCAGAGATGACAGACTGT 472
    |||

RESULT 12
US-10-066-543-1843
; Sequence 1843, Application US/10066543
; Publication No. US20030087818A1
; GENERAL INFORMATION:
; APPLICANT: Jjiang, Yugu
; APPLICANT: Xu, Jiangchun
; APPLICANT: Pyle, Ruth A.
; APPLICANT: Indritas, Carol Yoseph
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Fanger, Gary R.
; APPLICANT: Smith, Carole L.
; APPLICANT: Durham, Margarita
; APPLICANT: Stolk, John A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.563
; CURRENT APPLICATION NUMBER: US/10/066,543
; CURRENT FILING DATE: 2002-01-31
; NUMBER OF SEQ ID NOS: 3417
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1843
; LENGTH: 447
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-066-543-1843

Query Match      13.2%; Score 214.8; DB 14; Length 447;
Best Local Similarity 73.0%; Pred. No. 2.9e-54;
Matches 276; Conservative 0; Mismatches 102; Indels 0; Gaps 0;

QY 159 GCCCTTCTGATAGGGGTGAGCGCGGACATGCGCAGCGGAAAGTCCAGCGTGTGAGAA 218
    |||

```


Db 70 GCCCTTCCTTAATAGCGCTCAGCGGGGAAACAGTACGGGCAAGTCTTCCGTGTGCTAA 129
QY 219 GATCATGAGTGTGCTGGAGACAGAGTGGAAACAGCGGAGCGGAAAGTGTCTATCT 278
Db 130 GATGTGACAGTCTGTGGGAGAAATGAGTGGACTATTCGCAAGAGAGGTGTCTATCT 189
QY 279 GAGCCAGAGAGTGTCTAACAAGTCTGACGGCAGAGACAGAAAGGCCAAAGCCTTGAAG 338
Db 190 GAGCCAGAGATAGTCTTACCTGTCTTACCTCGAGACAGAAAGGCCAAAGCCTTGAAG 249
QY 339 ACAGTCAATTTTGGACATCCAGATGCTTGTATTAATGATTGACAGAGACTTGA 398
Db 250 CCAATTCACTTGAACCAACCGGATGCTTGAATGAATCAATCTCAAAAACATCA 309
QY 399 GAACATGTGAGGCGAAACAGTGGAGTGGCCACCTATGATTTTGTGACACTGAA 458
Db 310 AGAATACTGAAAGGAAACAGTCCAGATCCCGTATGACTTGTCTCCCAATTC 369
QY 459 GTTACAGAGACCAAGGTGTCTACCTGCGGAGCGTGTCTGTGAGGCGATCTTGT 518
Db 370 GAAGGAGAGACAGTACTGTCTATTCGCGACAGCTGTGTCTTTGAAGAACTCTGCC 429
QY 519 GTTCTACAGCCAGAGAT 536
Db 430 CTCTACTCCAGAGAT 447

RESULT 13

US-09-918-995-23923
; Sequence 23923, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 23923
; LENGTH: 455
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(455)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-23923

Query Match 13.0%; Score 210.4; DB 11; Length 455;
Best Local Similarity 72.8%; Pred. No. 6.3e-53;
Matches 272; Conservative 0; Mismatches 101; Indels 0; Gaps 0;

QY 159 GCCCTTCCTTAATAGCGGCTGACCGGCACTGCCAGGCGAAAGTGCACCTGTGTGAA 218
Db 84 GCCCTTCCTTAATAGCGGCTGACCGGCGAAGCAAGTCTTCCGTGTGTCTAA 143
QY 219 GATCATGAGTGTGCTGGAGACAGAGTGGAAACAGCGGAGCGGAAAGTGTCTATCT 278
Db 144 GATGTGACAGTCTGTGGGAGAAATGAGTGGACTATTCGCAAGAGAGGTGTCTATCT 203
QY 279 GAGCCAGAGAGTGTCTAACAAGTCTGACGGCAGAGACAGAAAGGCCAAAGCCTTGAAG 338
Db 204 GAGCCAGAGATAGTCTTACCGGTCTTACCTCGAGACAGAAAGGCCAAAGCCTTGAAG 263
QY 339 ACAGTCAATTTTGGACATCCAGATGCTTGTATTAATGATTGACAGAGACTTGA 398
Db 264 CCAATTCACTTGAACCAACCGGATGCTTGAATGAATCAATCTCAAAAACATCA 323
QY 399 GAACATGTGAGGCGAAACAGTGGAGTGGCCACCTATGATTTTGTGACACTGAA 458

Db 324 AGAATCAGTGAAGGAAACAGTCCAGATCCCGGTATGACTTTGTCTCCATTC 383
QY 459 GTTACAGAGACCAAGGTGTCTACCTTGGAGCGTGTCTTGTGAGGCGATCTTGT 518
Db 384 GAAGGAGAGACAGTACTGTCTATTCGCGACAGCTGTGTCTTGAAGGATCTTGGC 443
QY 519 GTTCTACAGCA 530
Db 444 CTCTACTCCCA 455

RESULT 14

US-10-029-386-15883
; Sequence 15883, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Hanzel, David R.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR GI
; FILE REFERENCE: A60MICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 15883
; LENGTH: 187
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR9.1
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.5
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.8
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.6
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.7
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.5
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.7
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.3
; OTHER INFORMATION: EST HUMAN HIT: A1992171.1, EVALUE 1.00e-100
; OTHER INFORMATION: SWISSPROT HIT: P52623, EVALUE 4.00e-26
; OTHER INFORMATION: NT HIT: g114783235, EVALUE 1.00e-100
US-10-029-386-15883

Query Match 11.4%; Score 185; DB 12; Length 187;
Best Local Similarity 100.0%; Pred. No. 1.9e-45;
Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 745 GTTGCATCAACTGATCGTGCAGACATCCAGAGACATTTGTGATGTGACATCTGCA 804
Db 3 GTTGCATCAACTGATCGTGCAGACATCCAGAGACATTTGTGATGTGACATCTGCA 62
QY 805 TGGCAGCGAGAGGCTCCATGCGCGAGCTCAACGCACTTTCTGAGCCAGGGGAC 864
Db 63 TGGCAGCGAGAGGCTCCATGCGCGAGCTCAACGCACTTTCTGAGCCAGGGGAC 122
QY 865 CACCTGAGATGTGACCTCTGCAACAGGTCAATTTGAGTCCAGACAGACCCAC 924
Db 123 CACCTGAGATGTGACCTCTGCAACAGGTCAATTTGAGTCCAGACAGACCCAC 182
QY 925 TGAGG 929
Db 183 TGAGG 187

RESULT 15

US-10-251-186-14
; Sequence 14, Application US/10251186
; Publication No. US20030180745A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom

```
/ APPLICANT: Zhang, Jie
/ APPLICANT: Dimanac, Radoje T.
/ TITLE OF INVENTION: No. US20030180745A1el Nucleic Acids and
/ TITLE OF INVENTION: Polypeptides
/ FILE REFERENCE: 789CIP2CDV1
/ CURRENT APPLICATION NUMBER: US/10/251,186
/ CURRENT FILING DATE: 2002-09-19
/ PRIOR APPLICATION NUMBER: 09/665,363
/ PRIOR FILING DATE: 2000-09-19
/ PRIOR APPLICATION NUMBER: 09/574,454
/ PRIOR FILING DATE: 2000-05-19
/ PRIOR APPLICATION NUMBER: 09/519,705
/ PRIOR FILING DATE: 2000-03-07
/ NUMBER OF SEQ ID NOS: 35
/ SOFTWARE: pc_fl_genes Version 2.0
/ SEQ ID NO 14
/ LENGTH: 2058
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (173)..(1909)
US-10-251-186-14
```

```
Query Match          10.8%; Score 175.6; DB 12; Length 2058;
Best Local Similarity 58.4%; Pred.No.4.9e-42;
Matches 326; Conservative 0; Mismatches 229; Indels 3; Gaps 1;

QY      236 GACGAAACGAGGTGGAAACAGCGGCGAAGCGGAAAGTGTGTCATCTTGAGCCGAGCAGGTTCT 295
Db      609 GAATGATCATGAGAGCCCTGATGTGCTCCCTGGGTGTCTTGTCTCCATGAGCTCCCTCT 668

QY      296 ACAAGTCTCTGACGCGCAGAGAGAGGCGCCCTGAAAGGACAGTACATTTTGACC 355
Db      669 ACAAGTCTCTGACGCGCAGAGAGAGGCGCCCGCACACACACTTCACTTGACC 728

QY      356 ATCCAGATGCCCTTTGATATGATTTGATGACAGAGACTCTGAAAGACATCGTGAAGGCA 415
Db      729 ACCCAGATGCCCTTTGATGATGATTTGATGACAGAGACTCTGAAAGGCAAGGGA 788

QY      416 AAACGGTGAAGTCCGACCTATGATTTTGTGACACACTCAAGTTTACAGAGACACCGG 475
Db      789 AGAGTGTCAAGTGCCTCCATTATGACTTCAACACGACAGCGGAGAGAGACTGGA AAA 848

QY      476 TGTCTACCCCTGCGGAGCGTGTGTTCTGTTGAGGGCATTTGTTCTACAGCCAGGAGA 535
Db      849 CACTGTATGTGCAAAAGCTCATCTTTGAGGGCATATGAGCTTTGCTGACAAAGACAC 908

QY      536 TCCGGGACATGTTCCACCTGCGCTCTTCTGTGACACCGACTCCGACGTGAGCTGTCTC 595
Db      909 TGTGAGAGCTCTGTGACATGAAGATCTTTGTGACACAGACTCCGACATCCGCTGGTAC 968

QY      596 GAAAGTTCTCCGGGAGCTGGG--CCGAGGAGAGGACCTGGAGCAGATTCTGACCGACT 652
Db      969 GCGGCTGCGCGCGGACATCGTGAAGCGCGGCGGACATGAGAGGTGTCAATCAAGCAGT 1028

QY      653 ACACCACTCTTGTAAGCGCGGCTTTGAGAGTTCGCTGCGGCAAAAGAAATATGCGG 712
Db      1029 ACAACAAGTTGTCAAGCTCTCTTTCGACGATACATCAAGCCCAACATGCGCTGGCAG 1088

QY      713 ATGTGATCATCCACGAGAGTGAACAATATGTTGCCATCAACCTGATCGTGCAGACA 772
Db      1089 ACATCGTGTCTCCCAAGAGGAGCGGCAACACGTTGGCCATGACCTGATTTGTGACGACG 1148

QY      773 TCCAGGACATTCTGAATG 790
Db      1149 TGACAGCCAGCTGGAGG 1166
```

Search completed: November 25, 2003, 02:18:49
Job time : 598.791 secs

PN WO200202761-A2

cDNA encoding novel
Human uridine kinase
DNA of a human urine
Human cDNA clone
Human mdt CDA In
Human transferrin
Human polynucleoti
Human prostate cat
Human ovarian anti
Drosophila melanog
Anticancer agent t
Human polynucleot
DNA encoding novel
Human gene express
Human prostate exp
Human prostate ex
Human prostate exp
Human prostate exp
Human prostate exp
Human prostate exp
CDNA encoding pro
Novel human coding
cDNA encoding nove
Drosophila melano
Drosophila melano
Drosophila melano
Oligonucleotide fd
Oligonucleotide fo
Zea mays DNA fragm
Human gene expres
Arabidopsis thalia
Human gene express
Aspergillus oryzae
Arabidopsis thalia

XX 10-JAN-2002.
PD 28-JUN-2001; 2001WO-US21063.
XX 30-JUN-2000; 2000US-216503P.
XX (MILL-) MILENNIUM PHARM INC.
XX Glucksmann MA;
XX WPI; 2002-140091/18.
DR P-PSDB; AAE16592.
XX
XX New isolated human uridine kinase family polypeptide 57658, useful for
PT treating hematopoietic neoplastic disorders and disorders of neurons,
PT heart and blood vessels -
XX
PS Claim 1; Fig 1a; 103pp; English.
XX
XX The patent discloses human uridine kinase-like polypeptides, designated
CC 57658 and polynucleotides encoding such proteins. 57658 DNAs and proteins
CC are useful for developing diagnostic and therapeutic agents for 57658-
CC mediated or related disorders such as hematopoietic neoplastic disorders
CC (e.g. leukaemia), haematopoietic disorders (e.g. psoriasis, dermatitis,
CC diabetes mellitus, rheumatoid arthritis, disorders of bone (e.g. rickets,
CC osteoporosis), heart (e.g. sarcoma, myocardial infarction), blood vessels
CC (e.g. hypertension, atherosclerosis) and neurons (e.g. Huntington's
CC disease, Alzheimer's disease, Parkinson's disease). 57658 sequences or
CC their antibodies are useful in screening assays, detection assays (e.g.
CC forensic biology) and predictive medicine (e.g. diagnostic assays,
CC prognostic assays, and monitoring clinical trials and pharmacogenomics).
CC They are useful as reagents for diagnosing and treating 57658-mediated
CC disorders. 57658 DNAs are useful in gene therapy, in chromosome mapping,
CC to identify an individual from a minute biological sample (tissue typing)
CC and to aid in forensic identification of the biological sample. The
CC present sequence is a DNA encoding human 57658 protein.
XX
SQ Sequence 1624 BP; 354 A; 427 C; 506 G; 337 T; 0 other:
Query Match 100.0%; Score 1624; DB 24; Length 1624;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1624; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GTGGGGTCTGCTTCGACCTCGGCGCTGGGCGGCGCGCGCGCGGGAAGGGCGGCG 60
DB 1 GTGGGGTCTGCTTCGACCTCGGCGCTGGGCGGCGCGCGCGCGGGAAGGGCGGCG 60
QY 61 GCGGGAGACCCGATGCGGGAGCGGAGCCGAGATGGCTTGGCGGGAGCGGAGACTGC 120
DB 61 GCGGGAGACCCGATGCGGGAGCGGAGCCGAGATGGCTTGGCGGGAGCGGAGACTGC 120
QY 121 GAGAGCCCGCGCGCGGAGGCGGACCGTCCGACCGAGCGGCGCTTCTGATAGGGGTAGC 180
DB 121 GAGAGCCCGCGCGCGGAGGCGGACCGTCCGACCGAGCGGCGCTTCTGATAGGGGTAGC 180
QY 121 GAGAGCCCGCGCGCGGAGGCGGACCGTCCGACCGAGCGGCGCTTCTGATAGGGGTAGC 180
DB 121 GAGAGCCCGCGCGCGGAGGCGGACCGTCCGACCGAGCGGCGCTTCTGATAGGGGTAGC 180
QY 181 GGGCGCACTGCGACCGGAGATCGACCGTGTGTGAGAAAGATCATGAGAGTTGCTGGGACAG 240
DB 181 GGGCGCACTGCGACCGGAGATCGACCGTGTGTGAGAAAGATCATGAGAGTTGCTGGGACAG 240
QY 241 AAGAGGTGGAACAGCGGCGAGCGGAGAGGTGTCATCTCGAGCGAGGACAGGTTTACAAG 300
DB 241 AAGAGGTGGAACAGCGGCGAGCGGAGAGGTGTCATCTCGAGCGAGGACAGGTTTACAAG 300
QY 301 GTCTCGACGCGAGACGAGAAAGCCCTTGAAGAGACAGTAAATTTTGAACATCCA 360
DB 301 GTCTCGACGCGAGACGAGAAAGCCCTTGAAGAGACAGTAAATTTTGAACATCCA 360
QY 361 GATCCCTTTGATTAATGATTTGATGACAGACTCTGAAGAACATCTGTGAGAGGCAAAACG 420
DB 361 GATCCCTTTGATTAATGATTTGATGATGACAGACTCTGAAGAACATCTGTGAGAGGCAAAACG 420
QY 421 GTGAGAGTGGCGACCTATGATTTTGTGACACACTCAAGGTTACAGAGACCAAGGTGTC 480

DB 421 GTGAGAGTGGCGACCTATGATTTTGTGACACACTCAAGGTTACAGAGACCAAGGTGTC 480
QY 481 TACCTTCGCGACGTGGTCTGTGTTGAGGGCATCTTGTTCTACAGCAGAGATCCG 540
DB 481 TACCTTCGCGACGTGGTCTGTGTTGAGGGCATCTTGTTCTACAGCAGAGATCCG 540
QY 541 GACATGTTCCACTTGGGCGCTTTCGTGACACCGAATCCGAGCTCAGGCTGTCCAGAA 600
DB 541 GACATGTTCCACTTGGGCGCTTTCGTGACACCGAATCCGAGCTCAGGCTGTCCAGAA 600
QY 601 GTTCTCCGAGACGTGGCGCGAGGAGGACCTTGACAGATTTCTGACGACATCACACC 660
DB 601 GTTCTCCGAGACGTGGCGCGAGGAGGACCTTGACAGATTTCTGACGACATCACACC 660
QY 661 TTCTGTAAGCGGCGCTTCAAGAGTTCCTGCTCCGACAAAGAAATATGCGATGATC 720
DB 661 TTCTGTAAGCGGCGCTTCAAGAGTTCCTGCTCCGACAAAGAAATATGCGATGATC 720
QY 721 ATCCACAGAGATGAGCAATATGTTGCTATCACTGATGCTGAGACATCCAGGAC 780
DB 721 ATCCACAGAGATGAGCAATATGTTGCTATCACTGATGCTGAGACATCCAGGAC 780
QY 781 ATTCTGAATGATGACATCTGCAAAATGACACCGAGAGGTTCCATAGGCGAGCTCAAG 840
DB 781 ATTCTGAATGATGACATCTGCAAAATGACACCGAGAGGTTCCATAGGCGAGCTCAAG 840
QY 841 CGGACCTTTTCTGAGCGAGGAGCAACCTGGAGTCTGACCTTGGCAACCGTCAAT 900
DB 841 CGGACCTTTTCTGAGCGAGGAGCAACCTGGAGTCTGACCTTGGCAACCGTCAAT 900
QY 901 TTGGAATCCAGACGACGACACCCCTGAGGGGCTGCGGACCTCAGGCGAGTCTCCGCG 960
DB 901 TTGGAATCCAGACGACGACACCCCTGAGGGGCTGCGGACCTCAGGCGAGTCTCCGCG 960
QY 961 CGGACATGTGTGTCAAGGAGCTGAGCTGGGAGCGCCACCAACCACTGCTTCTCTC 1020
DB 961 CGGACATGTGTGTCAAGGAGCTGAGCTGGGAGCGCCACCAACCACTGCTTCTCTC 1020
QY 1021 GGGCGACCCCGAGGAGGTGTATGACAGCGGCGCTTCTCACTGAGAGTGAATCAAG 1080
DB 1021 GGGCGACCCCGAGGAGGTGTATGACAGCGGCGCTTCTCACTGAGAGTGAATCAAG 1080
QY 1081 TGTGTCACTGACCTCAACTGCTGGGACATGACAGGGGTTCTGAGGTTTCAAGCCAC 1140
DB 1081 TGTGTCACTGACCTCAACTGCTGGGACATGACAGGGGTTCTGAGGTTTCAAGCCAC 1140
QY 1141 TTAGGCTCGTTGCGGTTTAAAGATCCCTCTAGGTCACTGAGAAATGCCAGAAATGTGCA 1200
DB 1141 TTAGGCTCGTTGCGGTTTAAAGATCCCTCTAGGTCACTGAGAAATGCCAGAAATGTGCA 1200
QY 1201 GGAAGCTTGGAGAGCTTCTGTGAGAAATGTGAGGACATTTATGGGAAATTTGAGAGAC 1260
DB 1201 GGAAGCTTGGAGAGCTTCTGTGAGAAATGTGAGGACATTTATGGGAAATTTGAGAGAC 1260
QY 1261 AGGCTAGACACTGCGCGGCTGATGTTTGTGACAGTAAACCAAGTGGGAGAGATT 1320
DB 1261 AGGCTAGACACTGCGCGGCTGATGTTTGTGACAGTAAACCAAGTGGGAGAGATT 1320
QY 1321 TTTTCCAGTGTGATCTGTTCTTACACACTCAACATTAACATAAATTTTGTGAA 1380
DB 1321 TTTTCCAGTGTGATCTGTTCTTACACACTCAACATTAACATAAATTTTGTGAA 1380
QY 1381 AGTACTTCTCTTTTAACTGATGATGTCCTCATGTTTCTGTTTCTGTTTCAATACA 1440
DB 1381 AGTACTTCTCTTTTAACTGATGATGTCCTCATGTTTCTGTTTCTGTTTCAATACA 1440
QY 1441 CAAGGCTGTTGTTGGCTTCAAACTTAATTTTCAATGACAGAGGTTTGAAGTCAAGCTG 1500
DB 1441 CAAGGCTGTTGTTGGCTTCAAACTTAATTTTCAATGACAGAGGTTTGAAGTCAAGCTG 1500
QY 1501 GCCTAACGGAATATGGGAGCGCACTGAGGAGATTTTCCCTTCTGCTTGAAG 1560

DB 1501 GCCTACACGGATATGGGGACCACTGAGGATGTTTCCCCCTTGCTTGCCTTAAG 1560
QY 1561 GCAGAGAAAGGAGGATGCTTGGAGAACCCAGATATACCCAGGCTTGTGGGGG 1620
DB 1561 GCAGAGAAAGGAGGATGCTTGGAGAACCCAGATATACCCAGGCTTGTGGGGG 1620
QY 1621 CCAG 1624
DB 1621 CCAG 1624

RESULT 2
AAH15621
ID AAH15621 standard; cDNA; 2160 BP.
XX
AC AAH15621;
DT 26-JUN-2001 (first entry)
XX
DE Human cDNA sequence SEQ ID NO:13951.
XX
KM Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000BP-0116126.
XX
PR 29-JUL-1999; 99JP-0248036.
PR 27-AUG-1999; 99JP-0300253.
PR 11-JAN-2000; 2000JP-0118776.
PR 02-MAY-2000; 2000JP-0183767.
PR 09-JUN-2000; 2000JP-0241899.
XX
XX
PA (HELI-) HELIX RES INST.
PI Oca T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
DR MPI; 2001-318749/34.
XX
PT Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs -
XX
XX
PS Claim 8; SEQ ID 13951; 2537pp + CD ROM; English.
XX
CC The present invention describes primer sets for synthesizing 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to a 5'-end
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB93893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.

XX
SQ Sequence 2160 BP; 457 A; 591 C; 671 G; 441 T; 0 other;
Query Match 98.5%; Score 1599.4; DB 22; Length 2160;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1614; Conservative 0; Mismatches 1; Indels 4; Gaps 1;
QY 6 GTGCTCTCCGACCTCGGCGCTGGGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 65
DB 7 GTGCTCTCCGACCTCGGCGCTGGGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 66
QY 66 GCCCGATGCGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCG 125
DB 67 GACCCGATGCGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCG 126
QY 126 CCCCGCGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCG 185
DB 127 CCCCGCGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCG 186
QY 186 CACTGCGACGGGAGGAGTGCACCGTGTGTGAGAGATCATGTGAGTGTGAGAGTGTGAGAG 245
DB 187 CACTGCGACGGGAGGAGTGCACCGTGTGTGAGAGATCATGTGAGTGTGAGAGTGTGAGAG 246
QY 246 GGTGGAACAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 305
DB 247 GGTGGAACAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 306
QY 306 GACGGGAGGAGCAAGAGGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCG 365
DB 307 GACGGGAGGAGCAAGAGGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCG 366
QY 366 CTTTGAATATGATTTGATGACAGAGCTCTGAAGAAATCTGTGAGAGGCGGAGGCGGAGG 425
DB 367 CTTTGAATATGATTTGATGACAGAGCTCTGAAGAAATCTGTGAGAGGCGGAGGCGGAGG 426
QY 426 GGTGCGGACCTATGATTTGATGACAGAGCTCTGAAGAAATCTGTGAGAGGCGGAGGCGGAGG 485
DB 427 GGTGCGGACCTATGATTTGATGACAGAGCTCTGAAGAAATCTGTGAGAGGCGGAGGCGGAGG 486
QY 486 TGGGAGCGTGTGTTCTGTTTGAAGGCGATCTTGTTGTTTGAAGGCGATCTTGTTGTTTGA 545
DB 487 TGGGAGCGTGTGTTCTGTTTGAAGGCGATCTTGTTGTTTGAAGGCGATCTTGTTGTTTGA 546
QY 546 GTTCCACCTGCGGCTCTTGTGTCGACACCGATCTCCGACGTCGTCTCGAAGATTTCT 605
DB 547 GTTCCACCTGCGGCTCTTGTGTCGACACCGATCTCCGACGTCGTCTCGAAGATTTCT 606
QY 606 CCGGGAGCGTGTGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 665
DB 607 CCGGGAGCGTGTGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 666
QY 666 GAAGCGGCGCTTGAAGAGTCTGCTGCGGACCAAGAAATATGCGATGATCATCC 725
DB 667 GAAGCGGCGCTTGAAGAGTCTGCTGCGGACCAAGAAATATGCGATGATCATCC 726
QY 726 ACAGAGAGTGTGACATATGTGTTGTCATCACTGATCTGTGACGACATCTCAGAGATTCT 785
DB 727 ACAGAGAGTGTGACATATGTGTTGTCATCACTGATCTGTGACGACATCTCAGAGATTCT 786
QY 786 GAATGTGACATCTGCAATATGTCACCGAGAGAGGTCCTGAGGAGGAGGAGGAGGAGGAGGAG 845
DB 787 GAATGTGACATCTGCAATATGTCACCGAGAGAGGTCCTGAGGAGGAGGAGGAGGAGGAGGAG 846
QY 846 GTTGTCTGAGCGAGGAGCAACCTGTGAGTGTGACCTTGTGCAACCGTCACTATTGGA 905
DB 847 GTTGTCTGAGCGAGGAGCAACCTGTGAGTGTGACCTTGTGCAACCGTCACTATTGGA 906
QY 906 GTTCAGAGCAACACCCCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 965
DB 907 GTTCAGAGCAACACCCCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 966
QY 966 TGTGTGTTAGGAGACTGAGCTGTGGAGAGCGCCACCAACCAACCACTGCTTCTGCGCGC 1025

Dh 967 TGTGTGTTACAGGAGCTGAGCGCTGGGGAGCGCCACCAACCCACTGCTTCTCGGCGC 1026
Qy 1026 ACCCCAGGGAGTGTATGACGCGAGCGCTTCTCACTGAGAGTGGAACTCAGATGTGT 1085
Dh 1027 ACCCCAGGGAGTGTATGACGCGAGCGCTTCTCACTGAGAGTGGAACTCAGATGTGT 1086
Qy 1086 CACTCAGACTCAACTTGTGGGACACTGACGAGGCTTCGAGGTTTTCAGCCACTTAA 1145
Dh 1087 CACTCAGACTCAACTTGTGGGACACTGACGAGGCTTCGAGGTTTTCAGCCACTTAA 1146
Qy 1146 CTGCTGCGGTTTAAAGATCCCTCTAGAGTCACTGAGAAATGCCAGAAATGTGAGAG 1205
Dh 1147 CTGCTGCGGTTTAAAGATCCCTCTAGAGTCACTGAGAAATGCCAGAAATGTGAGAG 1206
Qy 1206 CTTGGGAGGCTTCTGTGAGGAATGTGAGGACATTTATTTGGGAAATTTGAGAGACAG 1265
Dh 1207 CTTGGGAGGCTTCTGTGAGGAATGTGAGGACATTTATTTGGGAAATTTGAGAGACAG 1266
Qy 1266 AGACACTGGGCTGAGCTGATGTTTGTGACAGTGAACCCACAGTGGAGAGAGTTTTC 1325
Dh 1267 AGACACTGGGCTGAGCTGATGTTTGTGACAGTGAACCCACAGTGGAGAGAGTTTTC 1326
Qy 1326 CAGTCTGATCTGTTCTTACACACTCACACATTAAGTTTGTGAACTAGTAC 1385
Dh 1327 CAGTCTGATCTGTTCTTACACACTCACACATTAAGTTTGTGAACTAGTAC 1382
Qy 1386 TTTCTTTTTCATGTTTACATGTCCTCATGTTTCTGTTTCTGTTTCTATACACAGG 1445
Dh 1383 TTTCTTTTTCATGTTTACATGTCCTCATGTTTCTGTTTCTGTTTCTATACACAGG 1442
Qy 1446 CTGCTGTTGGGCTTCAAACTTATTTATGACCCAGTGGTTTGGAGTCCAGCGGAGCTTA 1505
Dh 1443 CTGCTGTTGGGCTTCAAACTTATTTATGACCCAGTGGTTTGGAGTCCAGCGGAGCTTA 1502
Qy 1506 CACGGATATGAGGAGCACTGAGGAGATGTTTCCCTGCTTGTGCTTAAAGAGAGA 1565
Dh 1503 CACGGATATGAGGAGCACTGAGGAGATGTTTCCCTGCTTGTGCTTAAAGAGAGA 1562
Qy 1566 GAAAGGAGGCGGAGTCCCTGGAAGACACCCAGATCAACCCAGGCTTGTGGGGGCGAG 1624
Dh 1563 GAAAGGAGGCGGAGTCCCTGGAAGACACCCAGATCAACCCAGGCTTGTGGGGGCGAG 1621

RESULT 3
AA160444
ID AA160444 standard; cDNA; 2152 BP.
XX
AC AA160444;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 4433.
XX
KW Human; nootropic; immunosuppressant; cyostatic; gene therapy; cancer; peripheral nervous system; neuropathy; central nervous system; CNS; Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic; amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic; chemokinetic; thrombolytic; drug screening; arthritis; inflammation; leukaemia; ss.
XX
OS Homo sapiens.
XX
PN WO200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US34263.
XX
PR 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.

PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou F, Goodrich R, Drmanac RT;
XX
DR WPI; 2001-442253/47.
DR P-PSDB; AAM41288.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
XX
PS Claim 1; SEQ ID NO 4433; 10078bp; English.
XX
CC The invention relates to human nucleic acids (AA157798-AA161369) and
CC the encoded polypeptides (AAM36642-AAM42213) with nootropic,
CC immunosuppressant and cyostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening
CC assays for receptor activity, arthritis and inflammation, leukaemia and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX
SQ Sequence 2152 BP; 468 A; 581 C; 660 G; 442 T; 1 other;

Query Match 97.6%; Score 1585.8; DB 22; Length 2152;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1587; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 36 CGCCGGGCGCCGGGAAAGGGGCGGCGCGGAGACCCGATCGCGGAGCGGAGCGCGAGAT 95
Dh 1 CGCCGGGCGCCGGGAAAGGGGCGGCGCGGAGACCCGATCGCGGAGCGGAGCGCGAGAT 60
Qy 96 GGCCTTGGGGGAGGCGGAAAGACTGCGAGAGCCCCCGCGGAGGCCGATCCGACCA 155
Dh 61 GGCCTTGGGGGAGGCGGAAAGACTGCGAGAGCCCCCGCGGAGGCCGATCCGACCA 120
Qy 156 GGGGCCCTTCGTGATGAGGGGTGAGCGCGGCACTGCGACGGGAAAGTCGACCGTGTGA 215
Dh 121 GCGGCCCTTCGTGATGAGGGGTGAGCGCGGCACTGCGACGGGAAAGTCGACCGTGTGA 180
Qy 216 GAAGATCATGAGATTCTGGGACAGAACGAGGTGGAACGCGGACGCGAAGGTGTCTAT 275
Dh 181 GAAGATCATGAGATTCTGGGACAGAACGAGGTGGAACGCGGACGCGAAGGTGTCTAT 240
Qy 276 CTTGAGCCAGAGAGGTTCTTCAAGGTCTTGACGGGAGAGCGAGAGCCAGGCTTGA 335
Dh 241 CTTGAGCCAGAGAGGTTCTTCAAGGTCTTGACGGGAGAGCGAGAGCCAGGCTTGA 300
Qy 336 AGGACGTACAATTTTGAACATCCAGATGCTTTGTAATGATTGATGACAGAGCTCT 395
Dh 301 AGGACGTACAATTTTGAACATCCAGATGCTTTGTAATGATTGATGACAGAGCTCT 360
Qy 396 GAAGAATCATGTGAGAGGCAAAAAGGTGAGAGTGCAGACTATGATTTGTGACACTC 455
Dh 361 GAAGAATCATGTGAGAGGCAAAAAGGTGAGAGTGCAGACTATGATTTGTGACACTC 420
Qy 456 AAGTTACAGAGACACAGGTGTCTTACCTTGGGAGCGTGTCTTGTGAGGCAATCTT 515
Dh 421 AAGTTACAGAGACACAGGTGTCTTACCTTGGGAGCGTGTCTTGTGAGGCAATCTT 480

114 AGACTGCGAGAGACCCCGCGCGAGGCGGACCGCTCCGACCAACGCGGCCCTTCTGATAGG 173
115 AGACTGCGAGAGACCCCGCGCGAGGCGGACCGCTCCGACCAACGCGGCCCTTCTGATAGG 120
174 GGTGAGCGCGCGGCACTGCGACGCGGGAAGTCCAGCCGTGTGTGAGAAATCATGAGAGTTGCT 233
121 GGTGAGCGCGCGGCACTGCGACGCGGGAAGTCCAGCCGTGTGTGAGAAATCATGAGAGTTGCT 180
234 GGGACAGAAAGAGGTGGAACAGCGGCGACGGAAGGTGTGATCTGAGCCGACGAGCAGGTT 293
181 GGGACAGAAAGAGGTGGAACAGCGGCGACGGAAGGTGTGATCTGAGCCGACGAGCAGGTT 240
294 CTACAGAGTCTGTGACGCGACAGAGCAAGAGCCCTTGAAGAGACAGTAACTTTTGA 353
241 CTACAGAGTCTGTGACGCGACAGAGCAAGAGCCCTTGAAGAGACAGTAACTTTTGA 300
354 CCATCCAGATGCTTTTGAATTAATGATTTGATGACAGGCTCTGAAAGAACTCGTGGAGGG 413
301 CCATCCAGATGCTTTTGAATTAATGATTTGATGACAGGCTCTGAAAGAACTCGTGGAGGG 360
414 CAAAACGCTGAGAGGTGCGGACCTATGATTTTGTGACACACTCAAGTTACAGAGACAC 473
361 CAAAACGCTGAGAGGTGCGGACCTATGATTTTGTGACACACTCAAGTTACAGAGACAC 420
474 GGTGCTCTACCTCTGCGGACGTGTGTTCTGTTTGAAGGGACCTGTTGTTCTACACCGAGA 533
421 GGTGCTCTACCTCTGCGGACGTGTGTTCTGTTTGAAGGGACCTGTTGTTCTACACCGAGA 480
534 GATCCGGGACATGTTTCCACCTGCGGCTCTTGTGTGACACCCACCTCCGACGTCAGGCTGTC 593
481 GATCCGGGACATGTTTCCACCTGCGGCTCTTGTGTGACACCCACCTCCGACGTCAGGCTGTC 540
594 TCGAAGAGTTTCCGCGGACGTGCGCCGACGAGGAGGAGCCTGAGAGAGATTTGACGACGTA 653
541 TCGAAGAGTTTCCGCGGACGTGCGCCGACGAGGAGGAGCCTGAGAGAGATTTGACGACGTA 599
654 CACCACTTGTGTGAACCGGCTCTGAGAGAGTTGCTGCTGCGGACAAAAGATATGCCGA 713
600 CACCACTTGTGTGAACCGGCTCTGAGAGAGTTGCTGCTGCGGACAAAAGATATGCCGA 659
714 TGTATCATCCGACGAGAGGTGACAAATATGTTGCCATCAACTGATCCGACGACAT 773
660 TGTATCATCCGACGAGAGGTGACAAATATGTTGCCATCAACTGATCCGACGACAT 719
774 CCAAGACATTTCTGATGTGACATCTGCAAAATGACACCGAAGAGGTGCCAAATGGGCGGAG 833
720 CCAAGACATTTCTGATGTGACATCTGCAAAATGACACCGAAGAGGTGCCAAATGGGCGGAG 779
834 CTACAGAGGACCTTTTCTGAGCCAGGGGACCACTTGGAGTGTGACCTTGTGCAAAACG 893
780 -TACAAAGCGGACCTTTTCTGAGCCAGGGGACCACTTGGAGTGTGACCTTGTGCAAAACG 838
894 GTCAATTTGAGATCCGACGAGACAGACCCCACTGAGGGGCTGCGGACGCTCAGGGCAGGTC 953
839 GTCAATTTGAGATCCGACGAGACAGACCCCACTGAGGGGCTGCGGACGCTCAGGGCAGGTC 898
954 TCCCGCCCGGATGTGTGAGGAGCTGA-GCCTGGGGAAGCGCCACCCACACCCACCTGC 1012
899 TCCCGCCCGGATGTGTGAGGAGCTGA-GCCTGGGGAAGCGCCACCCACACCCACCTGC 958
1013 TTCTCTTCGGCGCACCCCGAGGGAGTGTAGCAGCGAGGCTTCTCACTCAGAGAGTGA 1072
959 TTCTCTTCGGCGCACCCCGAGGGAGTGTAGCAGCGAGGCTTCTCACTCAGAGAGTGA 1018
1073 AACTCAGATGTGTACTCAGACTCACTGCTGGGACAGTGAAGAGGGGTTCTGAGAGTTT 1132
1019 AACTCAGATGTGTACTCAGACTCACTGCTGGGACAGTGAAGAGGGGTTCTGAGAGTTT 1078
1133 TCAAGCATTAGGCTCGTTGGGTTTAAAGATCCCTCTAGGTCACTGAGAAATGCCACAG 1192
1079 TCAAGCATTAGGCTCGTTGGGTTTAAAGATCCCTCTAGGTCACTGAGAAATGCCACAG 1138
1193 AATGTGAGGAAGCGCTGGAGAGCTTCTGTGTGAGGAATGTGAGGCACTTATTTGGGAAATT 1252

1139 AATGTGAGGAAGCGCTGGAGAGCTTCTGTGAGGAATGTGAGGACACTTATTTGGGAAATT 1198
1253 GAGAGACACACCTTAGACACTGCGTGGCTGATGTTTGTGTTGAGAGGAAACCCACAGTGGG 1312
1199 GAGAGACACACCTTAGACACTGCGTGGCTGATGTTTGTGTTGAGAGGAAACCCACAGTGGG 1258
1313 AGAGAGTTTTCAGTGTGATCTGTTCTTACACACTCAGACACATACTCAAAAGTTT 1372
1259 AGAGAGTTTTCAGTGTGATCTGTTCTTACACACTCAGACACATACTCAAAAGTTT 1314
1373 TGTGAACAACTACTTTCCTTTTGA 1397
1315 TGTGAACAACTACTTTCCTTTTGA 1339
RESULT 5
AA158658
ID AA158658 standard; cDNA, 1288 BP.
XX
AC AA158658;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 861.
XX
KW Human; noctropic; immunosuppressant; cyostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; hemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemoclastic;
KW chemokine; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia; ss.
XX
OS Homo sapiens.
XX
FN WO20015312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000MO-US34263.
XX
PR 21-JAN-2000; 2000US-0488725.
XX
PR 25-APR-2000; 2000US-0552317.
XX
PR 09-JUL-2000; 2000US-0598042.
XX
PR 19-JUL-2000; 2000US-0620312.
XX
PR 03-AUG-2000; 2000US-0653450.
XX
PR 14-SEP-2000; 2000US-0662191.
XX
PR 19-OCT-2000; 2000US-0693036.
XX
PR 29-NOV-2000; 2000US-0727344.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Weinman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
DR WPI; 2001-442253/47.
XX
DR P-PDSB; AAM95502.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders
XX such as central nervous system injuries -
XX
PS Claim 1; SEQ ID NO 861; 10078bp; English.
XX
CC The invention relates to human nucleic acids (AA157798-AA161369) and
CC the encoded polypeptides (AAM38642-AAW42213) with noctropic,
CC immunosuppressant and cyostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the

PR	11.-SEP.-2000	2000US.-0229344
PR	01.-SEP.-2000	2000US.-0229345
PR	05.-SEP.-2000	2000US.-0229349
PR	05.-SEP.-2000	2000US.-0229351
PR	06.-SEP.-2000	2000US.-0230437
PR	06.-SEP.-2000	2000US.-0230438
PR	06.-SEP.-2000	2000US.-0231242
PR	08.-SEP.-2000	2000US.-0231243
PR	08.-SEP.-2000	2000US.-0231244
PR	08.-SEP.-2000	2000US.-0231413
PR	08.-SEP.-2000	2000US.-0231414
PR	08.-SEP.-2000	2000US.-0233080
PR	12.-SEP.-2000	2000US.-0233081
PR	12.-SEP.-2000	2000US.-0231968
PR	14.-SEP.-2000	2000US.-0232397
PR	14.-SEP.-2000	2000US.-0232398
PR	14.-SEP.-2000	2000US.-0233359
PR	14.-SEP.-2000	2000US.-0233400
PR	14.-SEP.-2000	2000US.-0233401
PR	14.-SEP.-2000	2000US.-0233063
PR	14.-SEP.-2000	2000US.-0233064
PR	14.-SEP.-2000	2000US.-0233065
PR	21.-SEP.-2000	2000US.-0234223
PR	21.-SEP.-2000	2000US.-0234274
PR	25.-SEP.-2000	2000US.-0234997
PR	25.-SEP.-2000	2000US.-0234998
PR	25.-SEP.-2000	2000US.-0235484
PR	26.-SEP.-2000	2000US.-0235834
PR	27.-SEP.-2000	2000US.-0235836
PR	29.-SEP.-2000	2000US.-0236367
PR	29.-SEP.-2000	2000US.-0236368
PR	29.-SEP.-2000	2000US.-0236369
PR	29.-SEP.-2000	2000US.-0236370
PR	02.-OCT.-2000	2000US.-0237037
PR	02.-OCT.-2000	2000US.-0237038
PR	02.-OCT.-2000	2000US.-0237039
PR	13.-OCT.-2000	2000US.-0237040
PR	13.-OCT.-2000	2000US.-0241887
PR	20.-OCT.-2000	2000US.-0241878
PR	20.-OCT.-2000	2000US.-0241879
PR	20.-OCT.-2000	2000US.-0241886
PR	01.-NOV.-2000	2000US.-0244617
PR	08.-NOV.-2000	2000US.-0244674
PR	08.-NOV.-2000	2000US.-0244675
PR	08.-NOV.-2000	2000US.-0244676
PR	08.-NOV.-2000	2000US.-0244677
PR	08.-NOV.-2000	2000US.-0244678
PR	08.-NOV.-2000	2000US.-0244679
PR	08.-NOV.-2000	2000US.-0244682
PR	08.-NOV.-2000	2000US.-0244683
PR	08.-NOV.-2000	2000US.-0244685
PR	08.-NOV.-2000	2000US.-0244686
PR	08.-NOV.-2000	2000US.-0244687
PR	17.-NOV.-2000	2000US.-0249212
PR	17.-NOV.-2000	2000US.-0249213
PR	17.-NOV.-2000	2000US.-0249214
PR	17.-NOV.-2000	2000US.-0249211
PR	17.-NOV.-2000	2000US.-0249212
PR	17.-NOV.-2000	2000US.-0249213

PR	17-NOV-2000;	2000US-0249214.
PR	17-NOV-2000;	2000US-0249215.
PR	17-NOV-2000;	2000US-0249216.
PR	17-NOV-2000;	2000US-0249217.
PR	17-NOV-2000;	2000US-0249218.
PR	17-NOV-2000;	2000US-0249244.
PR	17-NOV-2000;	2000US-0249245.
PR	17-NOV-2000;	2000US-0249264.
PR	17-NOV-2000;	2000US-0249265.
PR	17-NOV-2000;	2000US-0249297.
PR	17-NOV-2000;	2000US-0249299.
PR	01-DEC-2000;	2000US-0249300.
PR	01-DEC-2000;	2000US-0250391.
PR	05-DEC-2000;	2000US-0251030.
PR	05-DEC-2000;	2000US-0251988.
PR	05-DEC-2000;	2000US-0256719.
PR	08-DEC-2000;	2000US-0251479.
PR	08-DEC-2000;	2000US-0251856.
PR	08-DEC-2000;	2000US-0251868.
PR	08-DEC-2000;	2000US-0251869.
PR	08-DEC-2000;	2000US-0251989.
PR	11-DEC-2000;	2000US-0251990.
PR	11-DEC-2000;	2000US-0254097.
PR	05-JAN-2001;	2001US-0259678.
XX		
PA	(HUMA-) HUMAN GENOME SCI INC.	
XX		
P1	Rosen CA, Baraesh SC, Ruben SM;	
XX		
DR	WPI; 2001-451937/48.	
XX		
PT	Isolated polypeptide for treating, preventing and/or prognosing	
PT	disorders related to the musculoskeletal system including	
PT	musculoskeletal cancers and also for testing and detection e.g.	
PT	diagnosis -	
XX		
PS	Example 2; SEQ ID NO 2804; 781bp + Sequence Listing; English.	
CC	The invention relates to novel genes (AAU34669-AAU37666) and proteins	
CC	(ABB03087-ABB04109) associated with the musculoskeletal system useful	
CC	for preventing, treating or ameliorating medical conditions e.g. by	
CC	protein or gene therapy. The genes are isolated from a range of human	
CC	tissues disclosed in the specification. The nucleic acids, proteins,	
CC	antibodies and (anti)agonists are useful in the diagnosis, treatment	
CC	and prevention of: (a) cancer, e.g. breast and ovarian cancer and	
CC	other cancers of the adrenal gland, bone, bone marrow, breast,	
CC	gastrointestinal tract, liver, lung, or urogenital; (b) immune	
CC	disorders e.g. Addison's disease, allergies, autoimmune haemolytic	
CC	anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,	
CC	multiple sclerosis, rheumatoid arthritis and ulcerative colitis;	
CC	(c) cardiovascular disorders such as myocardial ischemias; (d) wound	
CC	healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy;	
CC	and (f) infectious diseases such as viral, bacterial, fungal and	
CC	parasitic infections.	
CC	Note: The sequence data for this patent did not form part of the	
CC	printed specification, but was obtained in electronic format directly	
CC	from WIPO at ftp.wipo.int/pub/published_pct_sequences.	
XX		
SEQ	Sequence 9732 BP; 2063 A; 2805 C; 2561 G; 2303 T; 0 other;	
Query Match	54.2%; Score 880; DB 22; Length 9732;	
Best Local Similarity	100.0%; Pred. No. 1,9e-217;	
Matches 880; Conservative	0; Mismatches 0; Indels 0; Gaps 0;	
OY	745 GTTGCAATCAACCTGATCGTGCAGACAATTCGAATGGATCATGTCAA	804
Db	3901 TTGECATCAACTGATCGTGCAGACAATTCGAATGGATCATGTCAA	3966
OY	805 TTGCCACGAGAGGGTCCAATGGCGGAGCTTAACGCGAGCCTTTTCTGACCAGGAGAC	864
Db	3961 TTGCCACGAGAGGGTCCAATGGCGGAGCTTAACGCGAGCCTTTTCTGACCAGGAGAC	4022

QY 865 CACCCTGGATGCTGACCTCTGGCAAAAGGTCACTTTGAGATGTCACAGACAGACCCAC 924
DB 4021 CACCCTGGATGCTGACCTCTGGCAAAAGGTCACTTTGAGATGTCACAGACAGACCCAC 4080
QY 925 TAGAGGGGCTGCCGAGCTCAAGGGCAGGTCTCCGCCCGGCAATGTGTTCAGGGACTGAG 984
DB 4081 TGAGGGGCTGCCGAGCTCAAGGGCAGGTCTCCGCCCGGCAATGTGTTCAGGGACTGAG 4140
QY 985 CTTGGGGGAGGCCCAACCCCAACCCCACTGCTTCCCTGGGCCCAACCCCAAGGAGTTAGC 1044
DB 4141 CTTGGGGAGGCCCAACCCCAACCCCACTGCTTCCCTGGGCCCAACCCCAAGGAGTTAGC 4200
QY 1045 AGGAGGCTTCTCTCACTCAAGAGTGAATCTAGATGTGTCACTCACTCAACTTCTCT 1104
DB 4201 AGGAGGCTTCTCTCACTCAAGAGTGAATCTAGATGTGTCACTCACTCAACTTCTCT 4260
QY 1105 GGGACACTGACAGGCGTTCTCTGAGGTTTTCAGCCACTTAAAGCTTTCGTTTAAAGAT 1164
DB 4261 GGGACACTGACAGGCGTTCTCTGAGGTTTTCAGCCACTTAAAGCTTTCGTTTAAAGAT 4320
QY 1165 CCCTTAAAGTCACTGAGAAATGCGAAGATGTCAGAGAAAGCTGGAGGCTTCTGTGAG 1224
DB 4321 CCCTTAAAGTCACTGAGAAATGCGAAGATGTCAGAGAAAGCTGGAGGCTTCTGTGAG 4380
QY 1225 GAATGTGAGCACAATTATTTGGGAAATTGAGAGACAGCTTAACTGAGCTGAGCTGAT 1284
DB 4381 GAATGTGAGCACAATTATTTGGGAAATTGAGAGACAGCTTAACTGAGCTGAGCTGAT 4440
QY 1285 GTTTTGTGACAGTGAACCAAGTGGAGAGAGATTTTTCAGTCTGATCTGGTTCTTA 1344
DB 4441 GTTTTGTGACAGTGAACCAAGTGGAGAGAGATTTTTCAGTCTGATCTGGTTCTTA 4500
QY 1345 CACACTCAGACACATACTGAAAAGTTTGTGAACAGTACTTCTTTTCACTTA 1404
DB 4501 CACACTCAGACACATACTGAAAAGTTTGTGAACAGTACTTCTTTTCACTTA 4560
QY 1405 CATGCTCTCATGTTTCTGTTTCTGTTTCATTAACAAGAGCTGTGTGACCTACAAC 1464
DB 4561 CATGCTCTCATGTTTCTGTTTCTGTTTCATTAACAAGAGCTGTGTGACCTACAAC 4620
QY 1465 CTAAATTCATGACCCAGTGTGTTGACAGTCCAGCGTGGCTTACACGATATGGGAGGCAC 1524
DB 4621 CTAAATTCATGACCCAGTGTGTTGACAGTCCAGCGTGGCTTACACGATATGGGAGGCAC 4680
QY 1525 TGAGGGAGTGTTCCTCCCTTCTGTGCTTAAAGCACAAGAGCGGATGCTT 1584
DB 4681 TGAGGGAGTGTTCCTCCCTTCTGTGCTTAAAGCACAAGAGCGGATGCTT 4740
QY 1585 GGAAGCACCAGATCACACCCAGGCTTGTGCGGGGCGAG 1624
DB 4741 GGAAGCACCAGATCACACCCAGGCTTGTGCGGGGCGAG 4780

RESULT 7
ABX59427
ID ABX59427 standard; cDNA; 9732 BP.
AC ABX59427;
XX
XX
DT 26-FEB-2003 (first entry)
XX
XX
DE cDNA encoding novel human musculoskeletal system antigen #1771.
XX
XX
KM Gene; srf; musculoskeletal system antigen; cancer; metastasis;
KM re-vascularization; thrombosis; arteriosclerosis; mineral content;
KM cardiovascular condition; wound; injury; burn; angiogenesis; ulcer;
KM post-operative tissue repair; limb regeneration; neuronal growth;
KM neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
KM AIDS-related complex; chondrocyte growth; bone regeneration;
KM periodontal regeneration; tissue transport; bone graft; skin aging;
KM keratinocyte growth; hair loss; melanocyte growth; cell proliferation;
KM cell growth; organ transplant; cell differentiation; body height;
KM weight; hair colour; eye colour; skin; percentage of adipose tissue;

KM pigmentation; cosmetic surgery; metabolism; biorhythm; circadian rhythm;
KM depression; tendency for violence; pain; reproductive capability;
KM hormone level; endocrine level; appetite; libido; memory; stress;
KM storage capability; fat content; lipid content; protein content;
KM carbohydrate content; vitamin content; cofactor content;
KM nutritional component.
XX
OS Homo sapiens.
XX
XX US2002147140-A1.
XX
XX 10-OCT-2002.
XX
XX 17-JAN-2001; 2001US-0764877.
XX
XX 31-JAN-2000; 2000US-179065P.
XX
XX 04-FEB-2000; 2000US-180628P.
XX
XX 28-JUN-2000; 2000US-214886P.
XX
XX 07-JUL-2000; 2000US-216647P.
XX
XX 07-JUL-2000; 2000US-216880P.
XX
XX 11-JUL-2000; 2000US-217487P.
XX
XX 11-JUL-2000; 2000US-217496P.
XX
XX 14-JUL-2000; 2000US-218290P.
XX
XX 26-JUL-2000; 2000US-220963P.
XX
XX 26-JUL-2000; 2000US-220964P.
XX
XX 14-AUG-2000; 2000US-224518P.
XX
XX 14-AUG-2000; 2000US-224519P.
XX
XX 14-AUG-2000; 2000US-225267P.
XX
XX 14-AUG-2000; 2000US-225268P.
XX
XX 14-AUG-2000; 2000US-225270P.
XX
XX 14-AUG-2000; 2000US-225447P.
XX
XX 14-AUG-2000; 2000US-225757P.
XX
XX 14-AUG-2000; 2000US-225758P.
XX
XX 22-AUG-2000; 2000US-226868P.
XX
XX 30-AUG-2000; 2000US-228924P.
XX
XX 01-SEP-2000; 2000US-229287P.
XX
XX 01-SEP-2000; 2000US-229343P.
XX
XX 01-SEP-2000; 2000US-229344P.
XX
XX 01-SEP-2000; 2000US-229345P.
XX
XX 05-SEP-2000; 2000US-229509P.
XX
XX 05-SEP-2000; 2000US-229513P.
XX
XX 08-SEP-2000; 2000US-231413P.
XX
XX 21-SEP-2000; 2000US-234223P.
XX
XX 21-SEP-2000; 2000US-234274P.
XX
XX 25-SEP-2000; 2000US-234997P.
XX
XX 27-SEP-2000; 2000US-235834P.
XX
XX 29-SEP-2000; 2000US-236374P.
XX
XX 29-SEP-2000; 2000US-236375P.
XX
XX 29-SEP-2000; 2000US-236376P.
XX
XX 29-SEP-2000; 2000US-236386P.
XX
XX 29-SEP-2000; 2000US-236389P.
XX
XX 29-SEP-2000; 2000US-236390P.
XX
XX 29-SEP-2000; 2000US-236391P.
XX
XX 29-SEP-2000; 2000US-236392P.
XX
XX 02-OCT-2000; 2000US-237037P.
XX
XX 02-OCT-2000; 2000US-237038P.
XX
XX 02-OCT-2000; 2000US-237039P.
XX
XX 02-OCT-2000; 2000US-237040P.
XX
XX 13-OCT-2000; 2000US-239935P.
XX
XX 20-OCT-2000; 2000US-240960P.
XX
XX 20-OCT-2000; 2000US-241785P.
XX
XX 20-OCT-2000; 2000US-241809P.
XX
XX 01-NOV-2000; 2000US-244617P.
XX
XX 17-NOV-2000; 2000US-249299P.
XX
XX 08-DEC-2000; 2000US-251866P.
XX
XX 08-DEC-2000; 2000US-251868P.
XX
XX 08-DEC-2000; 2000US-251869P.
XX
XX (ROSE/) ROSEN C A.
XX (RUBE/) RUBEN S M.
XX (BARA/) BARASH S C.
XX
XX Rosen CA, Ruben SM, Barash SC,
XX WPI; 2003-128199/12.

XX Isolated nucleic acid molecules encoding musculoskeletal system
PT associated polypeptides, useful for detecting disorders, e.g. cancer -
XX
PS Disclosure; SEQ ID NO 2804; 321pp; English.
XX
CC The invention describes an isolated nucleic acid molecule comprising a
CC sequence encoding musculoskeletal system associated polypeptides useful
CC for detecting disorders, e.g., cancer or cancer metastases, in animals
CC or humans. The nucleic acid stimulates re-vascularisation of ischemic
CC tissues associated with conditions such as thrombosis, arteriosclerosis,
CC and other cardiovascular conditions; treats wounds due to injuries,
CC burns, post-operative tissue repair, and ulcers; stimulates angiogenesis
CC and limb regeneration; stimulates neuronal growth; can treat and prevent
CC neuronal damage occurring in certain disorders or neurodegenerative
CC conditions, such as, Alzheimer's disease, Parkinson's disease, and
CC AIDS-related complex; stimulates chondrocyte growth, thus they can be
CC used to enhance bone and periodontal regeneration and aid in tissue
CC transports or bone grafts; prevents skin aging due to sunburn by
CC stimulating keratinocyte growth; prevents hair loss, since FGF family
CC members activate hair-forming cells and promotes melanocyte growth;
CC stimulates growth and differentiation of hematopoietic cells and bone
CC marrow cells when used in combination with other cytokines; maintains
CC organs before transplantation or for supporting cell culture of primary
CC tissues; induces tissue of mesodermal origin to differentiate in early
CC embryos; increases or decreases the differentiation or proliferation of
CC embryonic stem cells, besides, haematopoietic lineage; modulates
CC mammalian characteristics, such as, body height, weight, hair colour, eye
CC colour, skin, percentage of adipose tissue, pigmentation, size, and shape
CC (e.g., cosmetic surgery); modulates mammalian metabolism; changes
CC mammal's meal state or physical state by influencing biorythms,
CC circadian rhythms, depression, tendency for violence, tolerance for pain,
CC reproductive capabilities, hormonal or endocrine levels, appetite,
CC libido, memory, or stress; increases or decreases storage capabilities,
CC fat content, lipid, protein, carbohydrate, vitamins, minerals, cofactors
CC or other nutritional components. This sequence encodes a novel human
CC musculoskeletal system antigen.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from the US patent office at
CC ftp.segdata.uspto.gov/sequence.html?DocID=20020147140.
CC
XX
SQ Sequence 9732 BP; 2063 A; 2805 C; 2561 G; 2303 T; 0 other;
Query Match 54.2%; Score 880; DB 25; Length 9732;
Best Local Similarity 100.0%; Pred. No. 1.9e-217;
Matches 880; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 745 GTTGCCATCAACCTGATCGTGACACATCCAGACATTGTAATGTCATCTGCAAA 804
Db 3901 GTTGCCATCAACCTGATCGTGACACATCCAGACATTGTAATGTCATCTGCAAA 3960
QY 805 TGGCACCAGAGAGGCTCAATGCGGAGCTACAAAGCGACCTTTTGAGCCAGGGGAC 864
Db 3961 TGGCACCAGAGAGGCTCAATGCGGAGCTACAAAGCGACCTTTTGAGCCAGGGGAC 4020
QY 865 CACCTGAGATGCTGACCTTGCGAAA CGGTCACTTTGAGTCACAGACAGACCCAC 924
Db 4021 CACCTGAGATGCTGACCTTGCGAAA CGGTCACTTTGAGTCACAGACAGACCCAC 4080
QY 925 TGAGGGGCTCGGAGCTCAGGGGAGAGTCTCCGCGGCACTGTGTTCAGGGACTGAG 984
Db 4081 TGAAGGGCTCGGAGCTCAGGGGAGAGTCTCCGCGGCACTGTGTTCAGGGACTGAG 4140
QY 985 CTTGGGAGCGGCCACCAACCACTGCTTCTCTGGCGGCAACCCAGGGGAGGTAGC 1044
Db 4141 CTTGGGAGCGGCCACCAACCACTGCTTCTCTGGCGGCAACCCAGGGGAGGTAGC 4200
QY 1045 AGCGAGGCTTCTCTCACTCAGAGTGAATCTCAGATGTCTCACTCAGACTCACTTGTCT 1104
Db 4201 AGCGAGGCTTCTCTCACTCAGAGTGAATCTCAGATGTCTCACTCAGACTCACTTGTCT 4260
QY 1105 GGGACACTGACAGGCGTTCTGAGGTTTTCAGGCCACTTAGGCTGTTGGGTTAAAGAT 1164

Db 4261 GGGACACTGACAGGCGTTCTGAGGTTTTCAGGCCACTTAGGCTGTTGGGTTAAAGAT 4320
QY 1165 CCTCTAGGTCATGAGAAATGCCACAGATGCGAGGAAGCCTGGAGGCTCTGTGAG 1224
Db 4321 CCTCTAGGTCATGAGAAATGCCACAGATGCGAGGAAGCCTGGAGGCTCTGTGAG 4380
QY 1225 GAATGTGAGCACATTAATGGGGAATTTAGAGACAGCCTTAGACACTGGCTGGCTGAT 1284
Db 4381 GAATGTGAGCACATTAATGGGGAATTTAGAGACAGCCTTAGACACTGGCTGGCTGAT 4440
QY 1285 GTTTGTGTCAGTGAACCCACAGTGGAGAGAGTTTTCAGTCTGATCTGTTCTTA 1344
Db 4441 GTTTGTGTCAGTGAACCCACAGTGGAGAGAGTTTTCAGTCTGATCTGTTCTTA 4500
QY 1345 CACAGTCACACATTAATCAAAAGTTTGTGAACAAGTACTTCCCTTTTTCATGTTA 1404
Db 4501 CACAGTCACACATTAATCAAAAGTTTGTGAACAAGTACTTCCCTTTTTCATGTTA 4560
QY 1405 CATGCTCATGTTTCTGTTTCTGTTTCATTAACAAGGCTGTTGTGCTTACAAAC 1464
Db 4561 CATGCTCATGTTTCTGTTTCTGTTTCTGTTTCATTAACAAGGCTGTTGTGCTTACAAAC 4620
QY 1465 CTAAATTCATGACCCAGTGTGTCAGTCCAGCTGCTTACACGGATATGGGAGCCAC 1524
Db 4621 CTAAATTCATGACCCAGTGTGTCAGTCCAGCTGCTTACACGGATATGGGAGCCAC 4680
QY 1525 TGAAGGATGTTTCCCCCTGCTGTGCTTAAAGAGAGAGGAGCGGATGCCCCCT 1584
Db 4681 TGAAGGATGTTTCCCCCTGCTGTGCTTAAAGAGAGAGGAGCGGATGCCCCCT 4740
QY 1585 GGAAGCACCAGCATCACACCAGGCTTGTGCGGGCCAG 1624
Db 4741 GGAAGCACCAGCATCACACCAGGCTTGTGCGGGCCAG 4780
RESULT 8
AAL36440
ID AAL36440 standard; DNA; 19125 BP.
XX
AC AAL36440;
XX
DT 08-JAN-2002 (first entry)
XX
DE Human musculoskeletal system related polynucleotide SEQ ID NO 2805.
XX
KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KW anti-allergic; hepatotropic; antidiabetic; anti-inflammatory; anticancer;
KW vlnelary; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW neurological disease; infection; human; secreted protein;
KW musculoskeletal system; ds.
XX
OS Homo sapiens.
PN WO20015367-A1.
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01338.
XX
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205615.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.

PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217486.
PR 14-JUL-2000; 2000US-0218280.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235836.
PR 27-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.

PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 06-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0256719.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
PR XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX WFI; 2001-451937/48.
XX
DR
XX
XX
PT Isolated polypeptide for treating, preventing and/ or prognosing
PT disorders related to the musculoskeletal system including
PT musculoskeletal cancers and also for testing and detection e.g.
PT diagnosis -
XX
XX
PS Example 2; SEQ ID NO 2805; 781pp + Sequence Listing; English.
XX
CC The invention relates to novel genes (A134669-A137666) and proteins
CC (A134669-A137666) associated with the musculoskeletal system useful
CC for preventing, treating or ameliorating medical conditions e.g. by
CC protein or gene therapy. The genes are isolated from a range of human
CC tissues disclosed in the specification. The nucleic acids, proteins,
CC antibodies and (ant)agonists are useful in the diagnosis, treatment
CC and prevention of: (a) cancer, e.g. breast and ovarian cancer and
CC other cancers of the adrenal gland, bone, bone marrow, breast,

CC gastrointestinal tract, liver, lung, or urogenital; (b) immune
CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis;
CC (c) cardiovascular disorders such as myocardial ischaemias; (d) wound
CC healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy;
CC and (f) infectious diseases such as viral, bacterial, fungal and
CC parasitic infections.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pat_sequences.

XX Sequence 19125 BP; 4263 A; 5265 C; 5444 G; 4153 T; 0 other;

Query Match 54.2%; Score 880; DB 22; Length 19125;

Best Local Similarity 100.0%; Pred. No. 2.6e-217; Mismatches 0; Indels 0; Gaps 0;

Matches 880; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 745 GTTGCATCAACCTGATGTCAGACATCCAGACATTTGAAATGTCATCTGCATA 804
DB 5997 GTTGCATCAACCTGATGTCAGACATCCAGACATTTGAAATGTCATCTGCATA 6056
QY 805 TGGCAACGAGAGGGTCCATGCGCGAGCTTCAACGCGACCTTTTCTGAGCCAGGGAC 864
DB 6057 TGGCAACGAGAGGGTCCATGCGCGAGCTTCAACGCGACCTTTTCTGAGCCAGGGAC 6116
QY 865 CACCTGGGATGTGACCTCTGGCAAAAGTTCACATTTGAGTCCAGAGAGACCCAC 924
DB 6117 CACCTGGGATGTGACCTCTGGCAAAAGTTCACATTTGAGTCCAGAGAGACCCAC 6176
QY 925 TGAAGGGGTGCGAGGCTCAGAGGAGGTCTCCCGCCGCGAGTGTGTTTCAAGGACTGAG 984
DB 6177 TGAAGGGGTGCGAGGCTCAGAGGAGGTCTCCCGCCGCGAGTGTGTTTCAAGGACTGAG 6226
QY 985 CTTGGGACGCCCAACCCACACCACTGCTCTCTCGCGCAACCCAGGGAGTTAGC 1044
DB 6237 CTTGGGACGCCCAACCCACACCACTGCTCTCTCGCGCAACCCAGGGAGTTAGC 6296
QY 1045 AGCGAGGCTTCTCTCAGTCAAGAGTGAATGTCAGTCACTCAAGCTCACTTGGCT 1104
DB 6297 AGCGAGGCTTCTCTCAGTCAAGAGTGAATGTCAGTCACTCAAGCTCACTTGGCT 6356
QY 1105 GGGACACTGACAGGCGTTCTGAGGTTTTCAGCCACTTACGCTGCTGCGTTTAAAGAT 1164
DB 6357 GGGACACTGACAGGCGTTCTGAGGTTTTCAGCCACTTACGCTGCTGCGTTTAAAGAT 6416
QY 1165 CCTCTTAGTCACTGAGAAATGCGACAGATGTGAGAGAGCTGGAGAGCTTCTGTGAG 1224
DB 6417 CCTCTTAGTCACTGAGAAATGCGACAGATGTGAGAGAGCTGGAGAGCTTCTGTGAG 6476
QY 1225 GAAATGTGAGGACATTTATGGGAAATGAGAGACAGCCAGACAGTGGCTGGCTGAT 1284
DB 6477 GAAATGTGAGGACATTTATGGGAAATGAGAGACAGCCAGACAGTGGCTGGCTGAT 6536
QY 1285 GTTTTGTGACAGTGAACCCACAGTGGAGAGAGTTTTCAGTCTGATCTGTTCTTA 1344
DB 6537 GTTTTGTGACAGTGAACCCACAGTGGAGAGAGTTTTCAGTCTGATCTGTTCTTA 6596
QY 1345 CACACTCACACATACTAACTAAAAGTTTGTGAACAAGTACTTTCTTTTACATGTTA 1404
DB 6597 CACACTCACACATACTAACTAAAAGTTTGTGAACAAGTACTTTCTTTTACATGTTA 6656
QY 1405 CATGTCCATGATTTTCTGTTTCTGTTTCTAATACAAAGGCTGGTGGCTTCAACAAC 1464
DB 6657 CATGTCCATGATTTTCTGTTTCTGTTTCTAATACAAAGGCTGGTGGCTTCAACAAC 6716
QY 1465 CTAAATTCATGACCCAGTGTGTCAGTCCAGCGTGGCTTACACGATATGGGGAGCCAC 1524
DB 6717 CTAAATTCATGACCCAGTGTGTCAGTCCAGCGTGGCTTACACGATATGGGGAGCCAC 6776
QY 1525 TGAAGGATGTTTTCCCTTCTGCTTGTGCTTAAAGGACAGAAAGCGAGGCGATGCCCT 1584
DB 6777 TGAAGGATGTTTTCCCTTCTGCTTGTGCTTAAAGGACAGAAAGCGAGGCGATGCCCT 6836

QY 1585 GGAAGCACCCAGCATCACACCCAGGCTTGGCGGGCCAG 1624
DB 6837 GGAAGCACCCAGCATCACACCCAGGCTTGGCGGGCCAG 6876

RESULT 9

ABX59428

AC ABX59428; ID ABX59428 standard; cDNA; 19125 BP.

XX 26-FEB-2003 (first entry)

DE cDNA encoding novel human musculoskeletal system antigen #1772.

XX Gene; ss; musculoskeletal system antigen; cancer; metastasis;
XX re-vascularisation; thrombosis; arteriosclerosis; mineral content;
XX cardiovascular condition; wound; injury; burn; angiogenesis; ulcer;
XX post-operative tissue repair; limb regeneration; neuronal growth;
XX neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
XX AIDS-related complex; chondrocyte growth; bone regeneration;
XX periodontal regeneration; tissue transport; bone graft; skin aging;
XX keratinocyte growth; hair loss; melanocyte growth; cell proliferation;
XX cell growth; organ transplant; cell differentiation; body height;
XX weight; hair colour; eye colour; skin; percentage of adipose tissue;
XX pigmentation; cosmetic surgery; metabolism; biorhythm; cardiac rhythm;
XX depression; tendency for violence; pain; reproductive capability;
XX hormone level; endocrine level; appetite; libido; memory; stress;
XX storage capability; fat content; lipid content; protein content;
XX carbohydrate content; vitamin content; cofactor content;
XX nutritional component.

OS Homo sapiens.

XX US2002147140-A1.

PN 10-OCT-2002.

PD 17-JAN-2001; 2001US-0764877.

XX 31-JAN-2000; 2000US-179065P.
XX 04-FEB-2000; 2000US-180628P.
XX 28-JUN-2000; 2000US-214886P.
XX 07-JUL-2000; 2000US-216647P.
XX 07-JUL-2000; 2000US-216880P.
XX 11-JUL-2000; 2000US-217487P.
XX 11-JUL-2000; 2000US-217496P.
XX 14-JUL-2000; 2000US-218290P.
XX 26-JUL-2000; 2000US-220963P.
XX 26-JUL-2000; 2000US-220964P.
XX 14-AUG-2000; 2000US-224518P.
XX 14-AUG-2000; 2000US-224519P.
XX 14-AUG-2000; 2000US-225267P.
XX 14-AUG-2000; 2000US-225268P.
XX 14-AUG-2000; 2000US-225270P.
XX 14-AUG-2000; 2000US-225447P.
XX 14-AUG-2000; 2000US-225757P.
XX 14-AUG-2000; 2000US-225758P.
XX 22-AUG-2000; 2000US-226868P.
XX 30-AUG-2000; 2000US-228924P.
XX 01-SEP-2000; 2000US-229287P.
XX 01-SEP-2000; 2000US-229343P.
XX 01-SEP-2000; 2000US-229344P.
XX 01-SEP-2000; 2000US-229345P.
XX 05-SEP-2000; 2000US-229509P.
XX 05-SEP-2000; 2000US-229513P.
XX 08-SEP-2000; 2000US-231413P.
XX 21-SEP-2000; 2000US-234223P.
XX 21-SEP-2000; 2000US-234274P.
XX 25-SEP-2000; 2000US-234997P.
XX 27-SEP-2000; 2000US-235834P.
XX 29-SEP-2000; 2000US-236327P.

XX	29-SEP-2000;	2000US-236367P.
PR	29-SEP-2000;	2000US-236368P.
PR	29-SEP-2000;	2000US-236369P.
PR	29-SEP-2000;	2000US-236370P.
PR	02-OCT-2000;	2000US-236802P.
PR	02-OCT-2000;	2000US-237037P.
PR	02-OCT-2000;	2000US-237038P.
PR	02-OCT-2000;	2000US-237039P.
PR	02-OCT-2000;	2000US-237040P.
PR	13-OCT-2000;	2000US-239355P.
PR	20-OCT-2000;	2000US-240960P.
PR	20-OCT-2000;	2000US-241785P.
PR	20-OCT-2000;	2000US-241809P.
PR	01-NOV-2000;	2000US-244617P.
PR	17-NOV-2000;	2000US-249299P.
PR	08-DEC-2000;	2000US-251856P.
PR	08-DEC-2000;	2000US-251868P.
XX	08-DEC-2000;	2000US-251869P.
XX		
PA	(ROSE/)	ROSEN C A.
PA	(RUBE/)	RUBEN S M.
PA	(BARA/)	BARASH S C.
XX		
XX	Rosen CA, Ruben SM, Barash SC,	
PI		
XX	WPI, 2003-128159/12.	
DR		
XX		
PT	Isolated nucleic acid molecules encoding musculoskeletal system	
XX	associated polypeptides, useful for detecting disorders, e.g. cancer -	
XX	Disclosure; SEQ ID NO 2805; 321pp; English.	

The invention describes an isolated nucleic acid molecule comprising a sequence encoding musculoskeletal system associated polypeptides useful for detecting disorders, e.g., cancer or cancer metastases, in animals or humans. The nucleic acid: stimulates re-vascularisation of ischemic tissues associated with conditions such as thrombosis, arteriosclerosis, and other cardiovascular conditions; treats wounds due to injuries, burns, post-operative tissue repair, and ulcers; stimulates angiogenesis and limb regeneration; stimulates neuronal growth; can treat and prevent neuronal damage occurring in certain disorders or neurodegenerative conditions, such as, Alzheimer's disease, Parkinson's disease, and AIDS-related complex; stimulates chondrocyte growth, thus they can be used to enhance bone and periodontal regeneration and aid in tissue transports or bone grafts; prevents skin aging due to sunburn by stimulating keratinocyte growth, prevents hair loss, since FGF family members activate hair-forming cells and promotes melanocyte growth; stimulates growth and differentiation of hematopoietic cells and bone marrow cells when used in combination with other cytokines; maintains organs before transplantation or for supporting cell culture of primary tissues; induces tissue of mesodermal origin to differentiate in early embryonic; increases or decreases the differentiation or proliferation of mammalian stem cells, besides, haematopoietic lineage; modulates mammalian characteristics, such as, body height, weight, hair colour, eye colour, skin, percentage of adipose tissue, pigmentation, size, and shape (e.g., cosmetic surgery); modulates mammalian metabolism; changes mammal's metal state or physical state by influencing biorythms, circadian rhythms, depression, tendency for violence, tolerance for pain, reproductive capabilities, hormonal or endocrine levels, appetite, libido, memory, or stress; increases or decreases storage capabilities, fat content, lipid, protein, carbohydrate, vitamins, minerals, cofactors or other nutritional components. This sequence encodes a novel human musculoskeletal system antigen.

Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from the US patent office at <http://seqdata.uspto.gov/sequence.html?docID=20020147140>.

SQ Sequence 19125 BP; 4263 A; 5265 C; 5444 G; 4153 T; 0 other;

Query Match	54.2%	Score 880;	DB 25;	Length 19125;
Best Local Similarly	100.0%	Pred. No. 2.6e-217;		
Matches 880; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	745	GTTCGCATCAACCGTATGCGACACATCCAGGACATTTGATGGTGACATCTGGCAA	804
Db	5997	GTTCGCATCAACCGTATGCGACACATCCAGGACATTTGATGGTGACATCTGGCAA	6056
QY	805	TGGCACCAGGAGGGGTCCAAATGGCGGAGACTCAAGGSGACCTTTTCTGACCAGGGAC	864
Db	6057	TGGCACCAGGAGGGGTCCAAATGGCGGAGACTCAAGGSGACCTTTTCTGACCAGGGAC	6116
QY	865	CACCTTGGGATGCTGACCTCTGGCAAAACGGTCACTTTGGAGTCAGACAGACCCAC	924
Db	6117	CACCTTGGGATGCTGACCTCTGGCAAAACGGTCACTTTGGAGTCAGACAGACCCAC	6176
QY	925	TGAGGGGCTGCGGACCTCAGGGGAGGATCTCCGCGCGGACATGATGTTACGGACTGAG	984
Db	6177	TGAGGGGCTGCGGACCTCAGGGGAGGATCTCCGCGCGGACATGATGTTACGGACTGAG	6238
QY	985	CCTGGGAGCGCCACCCACACCCCACTGCTCTCGGCGCACCCAGGGAGTGTAGC	1044
Db	6237	CCTGGGAGCGCCACCCACACCCCACTGCTCTCGGCGCACCCAGGGAGTGTAGC	6236
QY	1045	AGCAGGCTTCTCTCACTCAGGAGTGGAACTCAATGTGTCACTCAACTCAACTTGCT	1104
Db	6297	AGCAGGCTTCTCTCACTCAGGAGTGGAACTCAATGTGTCACTCAACTCAACTTGCT	6356
QY	1105	GGGACACTGACAGGCGTCTCGAGGTTTTCAGCCACTTAAGGCTGTCGGGTTAAAGAT	1164
Db	6357	GGGACACTGACAGGCGTCTCGAGGTTTTCAGCCACTTAAGGCTGTCGGGTTAAAGAT	6416
QY	1165	CCCTCTAGGTCACTGAGAAATGCCCACAGAAATGTGCAGGAAAGCTTGGAGGCTTGAG	1224
Db	6417	CCCTCTAGGTCACTGAGAAATGCCCACAGAAATGTGCAGGAAAGCTTGGAGGCTTGAG	6476
QY	1225	GAATGTAGGCACTTATTTGGGGAAATTTGAGGAGACAGCTTACACTGGCTGGCTGAT	1284
Db	6477	GAATGTAGGCACTTATTTGGGGAAATTTGAGGAGACAGCTTACACTGGCTGGCTGAT	6536
QY	1285	GTTTTGTGACAGGAACCCACAGTGGGAGAGATTTTTCAGTCTGATCTGGTCTTA	1344
Db	6537	GTTTTGTGACAGGAACCCACAGTGGGAGAGATTTTTCAGTCTGATCTGGTCTTA	6586
QY	1345	CACACTCACACATAACTCAAAAGTTTGTGAACAAGTACTTTCCTTTTTCATGTTA	1404
Db	6597	CACACTCACACATAACTCAAAAGTTTGTGAACAAGTACTTTCCTTTTTCATGTTA	6656
QY	1405	CATGTCTCATGTTTTCTGTTTTCTGTTTTCTATTAACAAGGCTGTGTGGCTTACAAC	1464
Db	6657	CATGTCTCATGTTTTCTGTTTTCTGTTTTCTATTAACAAGGCTGTGTGGCTTACAAC	6716
QY	1465	CTAATTTTCATGACCCAGGTGTTTTCAGTCCAGCGGTGCTTACAGGATPATTGGGAGCCAC	1524
Db	6717	CTAATTTTCATGACCCAGGTGTTTTCAGTCCAGCGGTGCTTACAGGATPATTGGGAGCCAC	6776
QY	1525	TGAGGGAATGTTTTCCCTCTGCTTGTGCTTTAAAGCAGAGACGAGCGGATGCTCT	1584
Db	6777	TGAGGGAATGTTTTCCCTCTGCTTGTGCTTTAAAGCAGAGACGAGCGGATGCTCT	6836
QY	1585	GGAGCAACCCAGCATCAACCCAGGCTTTGTGCGGGCCAG	1624
Db	6837	GGAGCAACCCAGCATCAACCCAGGCTTTGTGCGGGCCAG	6876
RESULT 10			
AAH75355			
ID	AAH75355 standard; cDNA; 900 BP.		
XX	AAH75355;		
XX	AC		
XX	02-OCT-2001 (first entry)		
DE	Human uridine kinase encoding cDNA.		
XX	Human; uridine kinase; UK; ss.		

RESULT 10
AAH75355
ID AAH75355 standard; cDNA; 900 BP.
XX
AC AAH75355;

DT	02-OCT-2001 (first entry)
XX	
DE	Human uridine kinase encoding cDNA
XX	
KW	Human; uridine kinase; UK; ss.

```
XX Homo sapiens.
OS
XX Key
XX CDS
XX Location/Qualifiers
XX 8..838
XX /*tag= a
XX /product= "uridine kinase"
XX /note= "claimed in claim 1"
XX
XX CN1287172-A.
XX
XX 14-MAR-2001.
XX
XX 07-SEP-1999; 99CN-0118818.
XX
XX 07-SEP-1999; 99CN-0118818.
XX
XX (UYFU-) UNIV FUDAN.
XX
XX Yu L, Zhao Y, Zhang H;
XX
XX WPI; 2001-409529/44.
XX
XX P-PSDB; AAG64506.
XX
XX Human uridine kinase and its coding sequence, preparation and
XX application.
XX
XX Claim 1; Page 15(Disclosure); 20pp; Chinese.
XX
XX The invention relates to human uridine kinase (UK).
XX
XX Sequence 900 BP; 201 A; 237 C; 292 G; 170 T; 0 other;
XX
XX
XX Query Match 52.6%; Score 855; DB 22; Length 900;
XX Best Local Similarity 97.5%; Pred. No. 2e-211;
XX Matches 880; Conservative 0; Mismatches 20; Indels 3; Gaps 1;
XX
XX
XX 87 GGCCGAGATGCTTCGGCGGAGGCGAAGACTGCGAGAGCCCGCGCGGAGCGGACCG 146
XX 1 GGCCGAGATGCTTCGGCGGAGGCGAAG---GTGCGAGAGCCCGCGCGGAGCGAAGCCG 57
XX
XX 147 TCCGCAACCAAGCGCCCTTCTGATAGGGGTAGCGCGCGCACTGCCAGCGGAAATCGAC 206
XX 58 TCCGCAACCAAGCGCTTCTGATAGGGGTAGCGCGCGCACTGCCAGCGGAAATCGAC 117
XX
XX 207 CGTGTGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 266
XX 118 CGTGTGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 177
XX
XX 267 GGTGATCATCTGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 326
XX 178 GGTGATCATCTGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 237
XX
XX 327 GGCTTTGAAAGAGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 386
XX 238 GGCTTTGAAAGAGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 297
XX
XX 387 CAGGACTCTGAAAGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 446
XX 298 CAGGACTCTGAAAGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 357
XX
XX 447 GACACACTCAAGGTTTACAGAGACCAAGGATGATGATGATGATGATGATGATGATGATGATG 506
XX 358 GACACACTCAAGGTTTACAGAGACCAAGGATGATGATGATGATGATGATGATGATGATGATG 417
XX
XX 507 GGGGATCTTGGTGTGTTTACAGAGACCAAGGATGATGATGATGATGATGATGATGATGATG 566
XX 418 GGGGATCTTGGTGTGTTTACAGAGACCAAGGATGATGATGATGATGATGATGATGATGATG 477
XX
XX 567 GGAACACGATCTCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 626
XX 478 GGAACACGATCTCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 537
XX
```

```
Qy 627 GGAACCTGAGACGATTTCTGACGAGTACACCACTTCGTGAAGCCGCGCTTCGAGAGTT 686
Dy 538 GGAACCTGAGACGATTTCTGACGAGTACACCACTTCGTGAAGCCGCGCTTCGAGAGTT 597
Qy 687 CTGCTGCGGAGAAAGATATCCGATGATGATGATGATGATGATGATGATGATGATGATGATG 746
Dy 598 CTGCTGCGGAGAAAGATATCCGATGATGATGATGATGATGATGATGATGATGATGATGATG 657
Qy 747 TGGCATCAACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 806
Dy 658 TGGCATCAACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 717
Qy 807 GCAACCGAGAGGATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 866
Dy 718 GCAACCGAGAGGATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 777
Qy 867 CCTGGGATGCTGATCTTGGCAAAAGGATGATGATGATGATGATGATGATGATGATGATGATG 926
Dy 778 CCTGGGATGCTGATCTTGGCAAAAGGATGATGATGATGATGATGATGATGATGATGATGATG 837
Qy 927 AGGGGCTGCGGAGGATTCAGGAGGATGATGATGATGATGATGATGATGATGATGATGATG 986
Dy 838 AGGGGCTGCGGAGGATTCAGGAGGATGATGATGATGATGATGATGATGATGATGATGATG 897
Qy 987 TGG 989
Dy 898 TGG 900
```

```
RESULT 11
AAK98735
ID AAK98735 standard; DNA; 834 BP.
XX
AC AAK98735;
XX
DT 02-MAY-2002 (first entry)
XX
DE DNA of a human uridine kinase (UDK).
XX
XX Human: uridine kinase; diagnostic assay; mutation detection; UDK;
XX probe; chromosome localization study; tissue expression; gene therapy;
XX antibody; vaccine; human ovarian cancer; immunological disorder;
XX human colon carcinoma; immunogen; ds.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 1..780
XX /*tag= a
XX /partial
XX /note= "No stop codon"
XX
XX WO200172963-A2.
XX
XX 04-OCT-2001.
XX
XX 27-MAR-2001; 2001MO-US09663.
XX
XX 27-MAR-2000; 2000US-0536647.
XX
XX (SMIK ) SMITHKLINE BEECHAM CORP.
XX
XX Ho YS, Johnson RK;
XX
XX WPI; 2001-626259/72.
XX
XX P-PSDB; AAO14412.
XX
XX Novel human uridine kinase polypeptides useful for treating cancers,
XX and to identify agonists and antagonists of the polypeptide useful for
XX treating conditions associated with uridine kinase imbalance
XX Claim 5; Page 22-23; 31pp; English.
XX
```


CC The invention relates to newly identified human uridine kinase (UDK)
CC polypeptides and polynucleotides and methods for producing such
CC polypeptides by recombinant techniques. Also disclosed in the invention
CC are methods for utilizing uridine kinase polypeptides and polynucleotides
CC in diagnostic assays. The polynucleotides and polypeptides of the
CC invention may be used as diagnostic reagents by detecting mutations in an
CC associated gene. An array of oligonucleotide probes comprising the
CC uridine kinase polynucleotide sequence or fragments thereof can be
CC constructed to conduct efficient screening of genetic mutations, for
CC example. Detection of abnormally decreased or increased levels of
CC polypeptide or mRNA expression may also be used for diagnosing or
CC determining susceptibility of a subject to a disease of the invention.
CC The polynucleotide sequences of the invention can be used for chromosome
CC localization studies and tissue expression studies. The polypeptides of
CC the invention or fragments thereof may be used as immunogens to produce
CC antibodies. These antibodies may be employed to isolate or identify
CC clones expressing the polypeptide. The polypeptides and polynucleotides
CC of the invention can be used as a vaccine or in gene therapy to treat
CC diseases such as human ovarian cancer, human colon carcinomas, and
CC immunological disorders. This polynucleotide sequence represents the DNA
CC of a human uridine kinase of the invention.

XX Sequence 834 BP; 194 A; 219 C; 265 G; 156 T; 0 other;

Query Match 51.3%; Score 832.4; DB 23; Length 834;
Best Local Similarity 99.9%; Pred. No. 1.4e-205;
Matches 833; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 94 ATGGCTTCGGCGGAGGCGAAGACTGCGAGAGCCCGCGCGAGCGGACCGTCCGAC 153
DB 1 ATGGCTTCGGCGGAGGCGAAGACTGCGAGAGCCCGCGCGAGCGGACCGTCCGAC 60
QY 154 CAGCGGCCCTCTCTGATGAGGGGTGAGCGCGGCACTCCGACCGGAAAGTGCACCTGTGT 213
DB 61 CACGCGGCCCTCTCTGATGAGGGGTGAGCGCGGCACTCCGACCGGAAAGTGCACCTGTGT 120
QY 214 GAGAGATCATGAGGTGCTGGGACAGACGAGGTGGAACAGGGCGCGGAAAGTGTGTC 273
DB 121 GAGAGATCATGAGGTGCTGGGACAGACGAGGTGGAACAGGGCGCGGAAAGTGTGTC 180
QY 274 ATCTGAGCCAGGACAGGTTCTACAGAGTCTTGAACGCGACAGACGAAAGGCGCTTG 333
DB 181 ATCTGAGCCAGGACAGGTTCTACAGAGTCTTGAACGCGACAGACGAAAGGCGCTTG 240
QY 334 AAAGGACAGTACAAATTTTGAACATCCAGATGCTTGTATATGATTTGATGACAGGACT 353
DB 241 AAAGGACAGTACAAATTTTGAACATCCAGATGCTTGTATATGATTTGATGACAGGACT 300
QY 394 CTGAGAGCATCGTGGAGGGCAAAAGGAGGAGGTGCGGACCTATGATTTTGTGACACAC 453
DB 301 CTGAGAGCATCGTGGAGGGCAAAAGGAGGAGGTGCGGACCTATGATTTTGTGACACAC 360
QY 454 TCAAGGTTACAGAGACCAACGAGTGTCTACCTGCGGACGTTGTTGTTGAGGCGATC 513
DB 361 TCAAGGTTACAGAGACCAACGAGTGTCTACCTGCGGACGTTGTTGTTGAGGCGATC 420
QY 514 TTGGTGTCTTACACGCGACGAGATCCGGGACATGTTCCACTGCGCTCTTCTGTGACACC 573
DB 421 TTGGTGTCTTACACGCGACGAGATCCGGGACATGTTCCACTGCGCTCTTCTGTGACACC 480
QY 574 GACTCCGAGTCAGGCGTGTCTCGAAGAGTCTCCGGGACGTGGGCGAGGAGGAGGACCTG 633
DB 481 GACTCCGAGTCAGGCGTGTCTCGAAGAGTCTCCGGGACGTGGGCGAGGAGGAGGACCTG 540
QY 634 GAGCAGATTTCTGACGACGATACCAACCTTGTGTAAGCGGCGCTTCTGAGAGATTCTGCTG 693
DB 541 GAGCAGATTTCTGACGACGATACCAACCTTGTGTAAGCGGCGCTTCTGAGAGATTCTGCTG 600
QY 694 CCGACAAAGAGATGCGGATGATATCCACAGAGAGTGAACAATATGTTGCCATC 753
DB 601 CCGACAAAGAGATGCGGATGATATCCACAGAGAGTGAACAATATGTTGCCATC 660
QY 754 AACCTGATCGTGAGACATCCAGGACATCTGAAATGTTGTAATCTGCAAAATGGACCGA 813

DB 661 AACCTGATCGTGAGACATCCAGGACATTTGAAATGTTGATCTGCAAAATGGACCGA 720
QY 814 GAGGAGTCCAAATGAGGCGAGCTACAAAGGACCTTTTCTGAGGACGAGGAGCAACCTGGG 873
DB 721 GAGGAGTCCAAATGAGGCGAGCTACAAAGGACCTTTTCTGAGGACGAGGAGCAACCTGGG 780
QY 874 ATCTGACCTCTGGCAAAAGGTTACATTTTGAATCCAGACGACGACCCCACTGA 927
DB 781 ATCTGACCTCTGGCAAAAGGTTACATTTTGAATCCAGACGACGACCCCACTGA 834

RESULT 12

AAH04832
ID AAH04832 standard; cDNA, 753 BP.
XX
XX AAH04832;
AC
XX 26-JUN-2001 (first entry)
DT
XX Human cDNA clone (5'-primer) SEQ ID NO:1667.
DE
XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX Homo sapiens.
XX
XX EP1074617-A2.
XX
XX 07-FEB-2001.
PD
XX 28-JUL-2000; 2000BP-0116126.
PF
XX 29-JUL-1999; 99JP-0248036.
PR 27-AUG-1999; 99JP-0300253.
PR 11-JAN-2000; 2000JP-0118776.
PR 02-MAY-2000; 2000JP-0183767.
PR 09-JUN-2000; 2000JP-0241899.
XX
XX (HELI-) HELIX RES INST.
XX
XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Iehi S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
XX WPI; 2001-318749/34.
DR
XX
XX
XX Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs -
XX
XX
XX Claim 1; SEQ ID 1667; 2537pp + CD ROM; English.
XX
XX The present invention describes primer sets for synthesizing 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide which comprises a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialized methods. AAH03166 to AAH13628 and
CC AAH1633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB9893 represent human amino acid sequences; and AAH13629 to AAH1632
CC represent oligonucleotides, all of which are used in the exemplification


```

Db      186 AGATCCGGGACATGTTCCACTGGGCTCTTCTGGAGACACGACTCCGAGTCAAGCTGT 245
Qy      593 CTGGAAGAGTTCTCCGGGAGCTGCGCCGAGAGGAGGACTCTGGACAGATTCTGACGAGT 652
Db      246 CTCGAAGAGTTCTCCGGGAGCTGCGCCGAGAGGAGGACTCTGGACAGATTCTGACGAGT 305
Qy      653 ACACCCACTTCTGGAGACCGGCGCTTGGAGAGTTCTGCGCGCG-----686
Db      306 ACACACCTTCTGGAGACCGGCGCTTGGAGAGTTCTGCGCGCGAGAGAGCTTGAC 365
Qy      697 -----ACAAAGAATATGC 710
Db      366 AGGGAATGAGAGTCAAGCATTTAGCCATAGAGTGTGGATGAGAGGAACAAAGAAATGTC 425
Qy      711 CGATGTGATATCCACGAGAGTGGACAATATGTTGCCATCAACTGATCTGTGACGA 770
Db      426 CGATGTGATATCCACGAGAGTGGACAATATGTTGCCATCAACTGATCTGTGACGA 485
Qy      771 CATCCAGGACATTTGTAATGTGACATCTGCAATGCGACCGAGAGGTCGAATGGCG 830
Db      486 CATCCAGGACATTTGTAATGTGACATCTGCAATGCGACCGAGAGGTCGAATGGCG 545
Qy      831 GAGCTACAAAGCGACCTTTCTGAGCGAGGAGCAACCTGGAGTCTGACTGTGCA 890
Db      546 GAGCTACAAAGCGACCTTTCTGAGCGAGGAGCAACCTGGAGTCTGACTGTGCA 605
Qy      891 ACGGTCACTTTGAGTGTGACGACGAGACCC 920
Db      606 ACGGTACATTTGAGTGTGAGCAGCANNCCGTCC 635

```

RESULT 14

AAH23801 standard; cDNA; 1322 BP.

AAH23801;

31-JUL-2001 (first entry)

Human transferrase HTFS-1 cDNA, SEQ ID NO:43.

Human transferrase; HTFS; agonist; antagonist; cellular signalling;
 proliferation; cell proliferative disorder; immune disorder;
 atherosclerosis; hepatitis; psoriasis; cancer; tumour;
 inflammation; AIDS; Addison's disease; allergy; asthma; anaemia;
 cirrhosis; Crohn's disease; atopic dermatitis; diabetes mellitus;
 multiple sclerosis; rheumatoid arthritis; pancreatitis;
 systemic lupus erythematosus; thrombocytopenia; ulcerative colitis;
 haemodialysis; extracorporeal circulation; trauma; transgenic animal;
 gene therapy; drug screening; ss.

Homo sapiens.

Location/Qualifiers

Key CDS 250..1025

/tag= a

/product= "HTFS-1"

/function= "Transferrase"

MO200132888-A2.

10-MAY-2001.

02-NOV-2000; 2000MO-US30485.

04-NOV-1999; 99US-0163595.

(INCY-) INCYTE GENOMICS INC.
 Tang YT, Yue H, Hillman JL, Lal P, Bandman O, Patterson C,
 Shih LT, Azimzai Y, Lu DAM, Baughn MR;

DR WPI: 2001-328796/34.
 DR P-FSDB; AAB73494.
 XX Human transferrase polypeptides and polynucleotides useful for
 PT diagnosis, prevention and treatment of cell proliferative and immune
 PT system disorders and for identifying agonists and antagonists -
 XX
 XX Claim 5; Page 135-136; 157pp; English.

Sequences AAB73494-AAB73535 represent novel human transferrase proteins
 HTFS-1 to HTFS-42, and sequences AAH23801-AAH23842 represent cDNAs
 encoding them. The proteins play important roles in the regulation of
 cellular signalling and proliferation. The HTFS proteins are useful for
 screening compounds for their effectiveness as agonists or antagonists of
 transferrase activity, or for compounds that specifically bind to an HTFS
 protein or which modulates the activity of an HTFS protein.
 CC Pharmaceutical compositions comprising an HTFS protein, HTFS
 CC agonist or antagonist, or genetic construct encoding an HTFS
 CC protein are useful for treating a disease or condition associated
 CC with decreased or increased expression of functional HTFS. Disorders
 CC which may be treated using such compositions include cell proliferative
 CC disorders and immune disorders. For example, diseases which may be
 CC treated include atherosclerosis, hepatitis, psoriasis, cancers (including
 CC breast, bladder, bone marrow, brain and uterus cancer), inflammation,
 CC AIDS, Addison's disease, allergies, asthma, anaemia, cirrhosis, Crohn's
 CC disease, atopic dermatitis, diabetes mellitus, multiple sclerosis,
 CC rheumatoid arthritis, pancreatitis, systemic lupus erythematosus,
 CC thrombocytopenia, and ulcerative colitis. They may also be used to treat
 CC complications of cancer, haemodialysis, extracorporeal circulation,
 CC trauma and haematopoietic cancer, including lymphoma, leukaemia and
 CC myeloma. Polynucleotides encoding HTFS proteins are useful for creating
 CC transgenic animals to model human diseases, for diagnostic purposes and
 CC to generate hybridisation probes useful in mapping the naturally
 CC occurring genomic sequences. HTFS, and its catalytic or immunogenic
 CC fragments are useful for screening libraries of compounds in a variety of
 CC drug screening techniques. Antibodies which specifically bind HTFS may be
 CC used for the diagnosis of disorders associated with the expression of
 CC HTFS, or in assays to monitor patients being treated with HTFS or
 CC agonists, antagonists or inhibitors of HTFS. The present sequence
 CC represents an HTFS protein-encoding cDNA of the invention.

Sequence 1322 BP; 339 A; 359 C; 358 G; 266 T; 0 other;

Query Match 23.4%; Score 379.6; DB 22; Length 1322;

Best Local Similarity 68.8%; Pred. No. 4.3e-88;

Matches 552; Conservative 0; Mismatches 244; Indels 6; Gaps 2;

```

Qy      14 CGACCTCGGCGCTGGGCGGCGCGCGCGGGAAGGGGCGGCGGAGCCCGAT 73
Db      158 CGGCTCGGAAGAACGAGGAGATGCCAGCGCGGCGGCGGAGCGTGCCTTTGG 217
Qy      74 GCGCGGAGCGGAGCGGCGGAGATGCTTCGCGGAGCGCAAGACTGCGAGCGCCGCG 133
Db      218 CACAGCGACGCGGAGGAGGCGGCGCGCAACATGCGCGGGGACGAGCAGACCTCTG 277
Qy      134 CGAGGCGGACCGTCCGACCAAGC--GCCCTTCTGATAGGGGTGAGCGCGGCACTG 190
Db      278 AGAACCAACAGCAGCAGCCAGCGCGGAGCCCTTCTTAAAGCGTCACCGGGGAACAG 337
Qy      191 CCAGCGGGAAGTGCAGCGGTGTGGAAGATGATGAGTGTGCGGAGCAAGACGAGTGG 250
Db      338 CTNAGCGCAAGTCTTCGTGTGTGTCTTAAGATGTGTGAGCTCTTGAGGAGAAATGAGTGG 397
Qy      251 AACAGCGGAGCGGAGAGTGTATCTTCTGAGCAGGAGCAGGTTCTAACAAGTCTTACCG 310
Db      398 ACTATGCCAGAGAGCAGGTGTATCTTCTGAGCAGGATGCTTCTTACCTTCTTACCT 457
Qy      311 CAGAGCAAGAGCCCAAGGCTTGAAGAGCAGTACATTTTGAACCTTCAAGATGCTTTG 370
Db      458 CGAGCAAGAGGCAAGGCGCTGAAGGCGCAATTTCACCAACCGGATGCTTTG 517
Qy      371 ATATATTTGATGACAGGACTCTGAAGACATCTGTGAGGCGAAGACGATGAGAGTGC 430

```

```

Db 518 ACAATGAATCATCTCTCAAAAACCTCAAAAGAAATCACTGAAGGAGGAAAAAGTCCAGATCC 577
Qy 431 CGACCTATGATTTTGTGACACACTCAAGGTTACAGAGACCGAGTGTGTTACCCGTGGG 490
Db 578 CCGGTATGACCTTTGTCTCCATTTCCGGAAGAGAGACAGTACTGTCTATCCCGCAG 637
Qy 491 ACGTGTCTCTGTGAGGGCATCTTGTGTTCTACAGCCAGAGAGATCCGGGACATGTTC 550
Db 638 ACGTGTCTCTGTGAGGATCTGTGGCTTCTACTCCAGAGAGATGAGACCTGTTC 697
Qy 551 ACGTGTCTCTGTGAGACCCGATCTCGAGTGAAGCTGTCTGAAAGTGTTCGGG 610
Db 698 AGATGAACCTTTTGTGATACAGATGCGACACCCGGCTCTCAGCGAGATATTAAAGG 757
Qy 611 ACGTGTCTCTGTGAGGAGGACCTGAGAGATCTGACGAGATCAACCACTTCGTGA 667
Db 758 ACATCAGGAGAGAGAGAGAGGATCTTGAGCAGATTTTATCTCAGTACATTACGTTCTCA 817
Qy 668 AGCCGGCTCTGAGAGATTTGCTGCTGCCGACAAAGAGATGCGAGTGTGATCCAC 727
Db 818 AGCTGCTTGAAGAAATTTGCTTGCCTTGCACAAAGAGATGCTGATGTGATCATCCCTA 877
Qy 728 GAGAGTGAACATATGTTGCTCCATCACTGATGTGCTGACAGACATCCAGACATTTCTGA 787
Db 878 GAGGTGAGATATCTGTGCTGATCAACCTCATCTGTGACAGACATCCAGACATCCCTGA 937
Qy 788 ATGTGACATCTGCAATGCA 809
Db 938 ATGAGGGCCCTCCAAAGCGCA 959

```

RESULT 15
AA157850
AA157850 standard; cDNA; 1402 BP.

AA157850;
22-OCT-2001 (first entry)
Human polynucleotide SEQ ID NO 53.

Human; noctropic; immunosuppressant; cyrostatic; gene therapy; cancer;
peripheral nervous system; neuropathy; central nervous system; CNS;
Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
chemokine; thrombolytic; drug screening; arthritis; inflammation;
leukemia; ss.

Homo sapiens.
MO200153312-A1.
26-JUL-2001.
26-DEC-2000; 2000WO-US34263.
21-JAN-2000; 2000US-0488725.
25-APR-2000; 2000US-0552317.
09-JUL-2000; 2000US-0598042.
19-JUL-2000; 2000US-0620312.
03-AUG-2000; 2000US-0653450.
14-SEP-2000; 2000US-0662191.
19-OCT-2000; 2000US-0693036.
29-NOV-2000; 2000US-0727344.

(HYSE-) HYSEQ INC.
Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
Zhang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
Zhao QA, Zhou P, Goodrich R, Drmanac RT;
WPI; 2001-442253/47.
P-PSDB; AAM8694.

XX Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
XX
PS Claim 1; SEQ ID NO 53; 10078bp; English.

XX The invention relates to human nucleic acids (AA157798-AA161369) and
CC the encoded polypeptides (AAM38642-AAM42213) with noctropic,
CC immunosuppressant and cyrostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.

XX Sequence 1402 BP; 379 A; 370 C; 371 G; 282 T; 0 other;

Query Match 23.2%; Score 377; DB 22; Length 1402;
Best Local Similarity 69.0%; Pred. No. 2, 1e-87;
Matches 546; Conservative 0; Mismatches 240; Indels 5; Gaps 2;

```

Qy 22 GCGCTGGGCGGCGCGCGCGGGAAGGCGCGCGCGGGAACCGGAGCGCGGA 81
Db 211 GCGGAGGAGATCCGACGCGCGCGCGCGCGCGCGCGCGTCCGTCGACAGGACGCG 270
Qy 82 GCGGAGGCGCGAGATGCGTCCGCGGAGGCGAGATCGAGAGCCCGCGCGGAGGCC 141
Db 271 GAGAGGCGCGCGCGGAACTGCGCGGAGCAGCAGACAGACCTTCAGAAACCAAGCA 330
Qy 142 GACCGTCCGACGAGCGCGCGCTTCTGATGAGGAGTGAAGCGGACCTGCGAGGGAAG 201
Db 331 GCCCAAGCGCGCG--GAGCGCTTCTTATAGGCGTCAAGCGGGAACAGTACGCGCAAG 388
Qy 202 TCGACCGTGTGAGAAAGATCATGAGTGTGGAACAGAACGAGGTGAACCGCGAG 261
Db 389 TCTTCCGCTGTCTAAAGATCGGACGCTCTGCGGACAGAAATGAGTGAAGTATCCGCA 448
Qy 262 CGGAGGTGTCTCTGAGCGGAGGAGGAGTGTGAGAGTGTGAGAGTGTGAGAGG 321
Db 449 AAGCAGGTGTCTCTGAGCGGAGGAGTGTGAGAGTGTGAGAGTGTGAGAGG 508
Qy 322 GCCAAGGCTTGAAGAGACAGTCAATTTTGAACATTCAGATGCTTTGATTAATTTG 381
Db 509 GCCAAGGCTTGAAGAGGCGGAGTCAATTTTGAACATTCAGATGCTTTGATTAATTTG 568
Qy 382 ATGCAAGGATCTGTAAGAACATGAGGAGGGAACAGGTGAGGTGCGGACCTATGAT 441
Db 569 ATTCTCAAAACATCAAGAAATCACTGAAGGGAAGAAACAGTCCAGATCCCGGTATGAC 628
Qy 442 TTTGTGACACATCAAGGTTCACAGAGACCAAGGTGTCTACCTTCGAGAGGTTCG 501
Db 629 TTTGTCTCCATTTCCGGAAGGAGGAGGAGTGTCTATCTGTATCCGCAACGAGGTCTC 688
Qy 502 TTTGAGGAGCATCTGTGTTTCTACAGCGGAGGATCCGGAAGATTTCCAGCTTCGCTC 561
Db 689 TTTGAGGAGATCTGTGCTTCTTACTCCGAGAGGTACAGAGACTGTTCAGATGAAGCTT 748
Qy 562 TTTGTGACACCGAGTCCAGCGTCAAGGTGTCTGAAAGGTTCGCGGAGGTGCC-- 618
Db 749 TTTGTGATATCAGATCCGAGACACCGGCTCTCAGCAGATATTAAAGACATCAAGCAG 808
Qy 619 CGAGGAGGAGGAGTGTGAGAGATTTGAGCGAGTACACACATTCGTTGGAAGCGGCGCTTC 678
Db 809 AGAGGAGGAGATCTTGAAGAGATTTTATCTCAGTATCAATTCGTTCAAGCGCTGCTTT 868

```


This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 24, 2003, 16:46:08 ; Search time 6310.34 Seconds

(without alignments)
10528.307 Million cell updates/sec

Title: US-09-896-522-1

Perfect score: 1624

Sequence: 1 gtcggggtcgcctccgcaccc.....ccagcttctgctggggccag 1624

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl: *
1: gb_ba: *
2: gb_htg: *
3: gb_in: *
4: gb_om: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pl: *
9: gb_pr: *
10: gb_ro: *
11: gb_sts: *
12: gb_sy: *
13: gb_un: *
14: gb_vi: *
15: em_ba: *
16: em_fun: *
17: em_hum: *
18: em_in: *
19: em_mu: *
20: em_om: *
21: em_ov: *
22: em_ov: *
23: em_pat: *
24: em_ph: *
25: em_pl: *
26: em_ro: *
27: em_sts: *
28: em_un: *
29: em_vi: *
30: em_htg_hum: *
31: em_htg_inv: *
32: em_htg_other: *
33: em_htg_mus: *
34: em_htg_pin: *
35: em_htg_mam: *
36: em_htg_rod: *
37: em_htg_vrt: *
38: em_sy: *
39: em_hugo_hum: *
40: em_hugo_mus: *
41: em_hugo_other: *

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1624	100.0	1624	6	AX449217 Sequence
2	1599.4	98.5	2160	6	BD157613 Primer fo
3	1599.4	98.5	2160	6	AK022317 Homo sapi
4	1453.4	89.5	2228	9	AK057848 Homo sapi
5	1397.4	86.0	2072	9	BC015547 Homo sapi
6	1156.8	71.2	1395	9	AF125106 Homo sapi
7	1018.8	62.7	1022	9	AF237290 Homo sapi
8	880	54.2	147492	9	AL358781 Human DNA
9	834	51.4	834	6	AX449219 Sequence
10	832.4	51.3	834	6	AF254133 Sequence
11	710	43.7	753	6	BD146824 Primer fo
12	666.6	41.0	1959	10	BC025146 Mus muscu
13	623	38.4	1810	10	MUSURK1
14	597.4	36.8	952	10	MOSURK1A
15	594.6	36.6	254933	2	AC078885 Mus muscu
16	574	35.3	192336	2	AC118474 Mus muscu
17	574	35.3	211075	10	AC076974 Mus muscu
18	483.6	29.8	734	6	AX540411 Sequence
19	379.6	23.4	1209	9	BC002906 Homo sapi
20	379.6	23.4	1322	6	AX135546 Sequence
21	374	23.0	744	9	AB062451 Homo sapi
22	374	23.0	1121	9	AF236637 Homo sapi
23	364.6	22.5	1784	5	BC045968 dario rer
24	363.2	22.4	1312	10	BC023789 Mus muscu
25	361.6	22.3	1354	10	AF236636 Mus muscu
26	300.4	18.5	1467	3	AY119583 Drosophill
27	255.6	15.7	841	11	BV017303 S21P6359
28	182.2	11.2	831	9	D78335 Human mRNA
29	177.2	10.9	1847	6	AX463241 Sequence
30	177.2	10.9	1847	9	AK000524 Homo sapi
31	175.8	10.8	1875	9	BC033078 Homo sapi
32	175.6	10.8	1813	6	AX405842 Sequence
33	159.6	9.8	1823	10	BC016535 Mus muscu
34	155.6	9.6	156905	2	AC020006 Drosophill
35	155.6	9.6	199016	3	AC008204 Drosophill
36	155.6	9.6	206741	3	AB003747 Drosophill
37	153.2	9.4	314	10	AB030700 Rattus no
38	150.2	9.2	336	9	BT006860 Homo sapi
39	150.2	9.2	336	12	BT007548 Synthetic
40	146.2	9.0	233210	2	AC097693 Rattus no
41	146.2	9.0	244105	2	AC098897 Rattus no
42	145.2	8.9	3246	3	AY129436 Drosophill
43	140.8	8.7	241882	10	AL808027 Mouse DNA
44	140.8	8.7	254677	2	AC068494 Mus muscu
45	136.2	8.4	655	11	BV077340 S21P6009

ALIGNMENTS

RESULT 1
AX449217
LOCUS AX449217 1624 bp DNA linear PAT 03-JUL-2002
DEFINITION Sequence 1 from Patent WO0202761.
ACCESSION AX449217
VERSION AX449217.1 GI:21697994
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Glucksmann, M. A.
TITLE 57658, a human uridine kinase and uses thereof
JOURNAL Patent: WO 0202761-A 1 10-JAN-2002;

Pred. No. is the number of results predicted by chance to have a

Millennium Pharmaceuticals, Inc. (US)
Location/Qualifiers
1. 1624
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
94..927
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD37771.1"
/db_xref="GI:21697995"
/translation="MASAGSDPSPAPADRPHPFLIGVSGTASGKSTCEKIM
ELIGNEVEBQRKVILISODRKYUTLAKOKAKLGOVNPDPHAPAFNDLHRLTK
NIVEGKVEVPTIDFVTHSRUPETTVYRPADVLPFEBILVPSOEIYDMNHLPLPDT
DSIVLRRLRVLRGRDLEQLITQYTFVKPAFEFCLEPTKKYADVILPRVDNNV
AIIIVIOHIDILNGDICKWHRGSGNSRYKRTFSEBPDHMLTSGKRLHSSRP
H"
BASE COUNT 354 a 427 c 506 g 337 t
ORIGIN
Query Match 100.0%; Score 1624; DB 6; Length 1624;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1624; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GTGGGTCGCTCGACCTGGCGGCTGGCGGGCGCGCGCGGGGAAAGGGCGGGC 60
DB 1 GTGGGTCGCTCGACCTGGCGGCTGGCGGGCGCGCGCGGGGAAAGGGCGGGC 60
QY 61 GCGGGAGCCCGATCGCGGAGCGGAGCCGAGATGCTTCGCGGGAGCGAAGCTGC 120
DB 61 GCGGGAGCCCGATCGCGGAGCGGAGCCGAGATGCTTCGCGGGAGCGAAGCTGC 120
QY 121 GAGAGCCCCGCGCGGAGCGGAGCCGCTCCGACCAAGCGGCTTCGTATGSGGTAGC 180
DB 121 GAGAGCCCCGCGCGGAGCGGAGCCGCTCCGACCAAGCGGCTTCGTATGSGGTAGC 180
QY 121 GAGAGCCCCGCGCGGAGCGGAGCCGCTCCGACCAAGCGGCTTCGTATGSGGTAGC 180
DB 121 GAGAGCCCCGCGCGGAGCGGAGCCGCTCCGACCAAGCGGCTTCGTATGSGGTAGC 180
QY 181 GCGCGCACTGCGACGGGGAAGTCACCGTGTGTGAAGAACATGAGAGTGTGGAGCAG 240
DB 181 GCGCGCACTGCGACGGGGAAGTCACCGTGTGTGAAGAACATGAGAGTGTGGAGCAG 240
QY 181 GCGCGCACTGCGACGGGGAAGTCACCGTGTGTGAAGAACATGAGAGTGTGGAGCAG 240
DB 181 GCGCGCACTGCGACGGGGAAGTCACCGTGTGTGAAGAACATGAGAGTGTGGAGCAG 240
QY 241 AACGAGGTGAAACAGCGGACGAGAAAGTGTCTCATCTGAGCCAGAGCAGGTTCTAACG 300
DB 241 AACGAGGTGAAACAGCGGACGAGAAAGTGTCTCATCTGAGCCAGAGCAGGTTCTAACG 300
QY 301 GTCTGACGCGACGAGCAGAGGCCAAGGCTTGAAGAGCAGTACAAATTTGACATCCA 360
DB 301 GTCTGACGCGACGAGCAGAGGCCAAGGCTTGAAGAGCAGTACAAATTTGACATCCA 360
QY 361 GATGCTTTGATTAATGATTGATGACAGCACTCAAGGTTACAGAGCAGGTTGTC 420
DB 361 GATGCTTTGATTAATGATTGATGACAGCACTCAAGGTTACAGAGCAGGTTGTC 420
QY 421 GTGAGGTGCGGACCTATGATTTTGTGACACACTCAAGGTTACAGAGCAGGTTGTC 480
DB 421 GTGAGGTGCGGACCTATGATTTTGTGACACACTCAAGGTTACAGAGCAGGTTGTC 480
QY 481 TACCTGCGGACGAGTGTCTGTTGAGGAGCATTTGGTGTCTACAGCAGAGATCCGG 540
DB 481 TACCTGCGGACGAGTGTCTGTTGAGGAGCATTTGGTGTCTACAGCAGAGATCCGG 540
QY 541 GACATGTTCCACTGCGGCTCTTGTGTGACACCGACTCCGACGCTGTCTCGAAGA 600
DB 541 GACATGTTCCACTGCGGCTCTTGTGTGACACCGACTCCGACGCTGTCTCGAAGA 600
QY 601 GTTCTCGGAGAGTGTGCGCGGAGGAGGAGCTTGAGAGCATTTGACGACAGTACACAC 660
DB 601 GTTCTCGGAGAGTGTGCGCGGAGGAGGAGGAGCTTGAGAGCATTTGACGACAGTACACAC 660
QY 661 TTGCTGAAGCGGCTTGTGAGAGATTTCTGCTGCGCAAGAAAGATGCGGATGTGATC 720
DB 661 TTGCTGAAGCGGCTTGTGAGAGATTTCTGCTGCGCAAGAAAGATGCGGATGTGATC 720
QY 721 ATCCACGAGAGAGTGAACAATATGTTGCCATCAACCTGATCGTGCAGCATCCAGAGC 780

DB 721 ATCCACGAGAGAGTGAACAATATGTTGCCATCAACCTGATGTGCAGCATCCAGAGC 780
QY 781 ATTCTGAATGTGTACATCTGCAATATGCAACGAGAGGTTCCAAATGGCGGAGTCAAG 840
DB 781 ATTCTGAATGTGTACATCTGCAATATGCAACGAGAGGTTCCAAATGGCGGAGTCAAG 840
QY 841 CGGACCTTTTCTGAGCGGAGGAGCCACCTTGGGATGCTGACCTTGGCAAAAGGTACAT 900
DB 841 CGGACCTTTTCTGAGCGGAGGAGCCACCTTGGGATGCTGACCTTGGCAAAAGGTACAT 900
QY 901 TTGAGTCCAGCAGACGAGACCCCACTGAGGAGGCTGCGGAGCTCAGGCGAGTCTCCGCC 960
DB 901 TTGAGTCCAGCAGACGAGACCCCACTGAGGAGGCTGCGGAGCTCAGGCGAGTCTCCGCC 960
QY 961 CGGACATGTGTGTTCAAGGAGCTGAGCTGGGAGCGCCACCCACACCTGCTCTCTC 1020
DB 961 CGGACATGTGTGTTCAAGGAGCTGAGCTGGGAGCGCCACCCACACCTGCTCTCTC 1020
QY 1021 GCGGACCCCCAGGAGAGTGTAGACAGGAGGCTTCTGACCTCAGAGAGTGGAACTCAGA 1080
DB 1021 GCGGACCCCCAGGAGAGTGTAGACAGGAGGCTTCTGACCTCAGAGAGTGGAACTCAGA 1080
QY 1081 TGTGTCACTCAGACTCAACTTGTGTGGAGACCTGACAGGAGTCTGAGGTTTTCAGCCAC 1140
DB 1081 TGTGTCACTCAGACTCAACTTGTGTGGAGACCTGACAGGAGTCTGAGGTTTTCAGCCAC 1140
QY 1141 TTAGGCTCGTTGGGTTTAAAGATCCCTTGAAGTCACTGAGAAATGCCACAGATGTGA 1200
DB 1141 TTAGGCTCGTTGGGTTTAAAGATCCCTTGAAGTCACTGAGAAATGCCACAGATGTGA 1200
QY 1201 GGAAGCTGGAGAGCTTCTGTAGGAAATGTAGAGGCACTTAATGGGGAATTAAGAGAGC 1260
DB 1201 GGAAGCTGGAGAGCTTCTGTAGGAAATGTAGAGGCACTTAATGGGGAATTAAGAGAGC 1260
QY 1261 AGCCTAGACACTGGCTGCTGATGTTTGTGACAGTGAACCCACAGTGGAGAGATT 1320
DB 1261 AGCCTAGACACTGGCTGCTGATGTTTGTGACAGTGAACCCACAGTGGAGAGATT 1320
QY 1321 TTTTCCAGTGTGATGCTGTTTCAACACTCACACATPACTCAAAAGTTTGTGAACA 1380
DB 1321 TTTTCCAGTGTGATGCTGTTTCAACACTCACACATPACTCAAAAGTTTGTGAACA 1380
QY 1381 AGTACTTCTCTTTTCAAGTATGATCAATGCTCAATGCTTCTGTTTCTGTTTCAATACA 1440
DB 1381 AGTACTTCTCTTTTCAAGTATGATCAATGCTCAATGCTTCTGTTTCTGTTTCAATACA 1440
QY 1441 CAAGGCTGTTGTGGCTCAAACTTAATTTCACTGACCGAGTGTGACAGTCCAGCTG 1500
DB 1441 CAAGGCTGTTGTGGCTCAAACTTAATTTCACTGACCGAGTGTGACAGTCCAGCTG 1500
QY 1501 GCTTACACGAGATATGAGGAGCCACTGAGGAGTGTTCCTGCTGCTGCTTAAAG 1560
DB 1501 GCTTACACGAGATATGAGGAGCCACTGAGGAGTGTTCCTGCTGCTGCTTAAAG 1560
QY 1561 GCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1620
DB 1561 GCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1620
QY 1621 CCAG 1624
DB 1621 CCAG 1624
RESULT 2
BD157613 2160 bp DNA linear PAT 17-JAN-2003
LOCUS BD157613
DEFINITION Primer for synthesizing full-length cDNA and use thereof.
ACCESSION BD157613
VERSION BD157613.1 GI:27863371
KEYWORDS JP 2002191363-A/12456.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

LOCUS AK022317 2160 bp mRNA linear PRI 01-AUG-2002
DEFINITION Homo sapiens CDNA FLJ12255 f18, clone MAMMA1001476, highly similar
to URIDINE KINASE (EC 2.7.1.48).
ACCESSION AK022317
VERSION AK022317.1 GI:10433687
KEYWORDS oligo capping; f18 (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Isegai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y.,
Nishikawa,T., Nagai,K., Sugano,S., Shiratori,A., Sudo,H.,
Magatsuma,M., Hosofiri,T., Kaku,Y., Kodaira,H., Kondo,H.,
Sugawara,M., Takahashi,M., Chiba,Y., Ishida,S., Murakawa,K.,
Ono,Y., Takiguchi,S., Watanabe,S., Kimura,K., Murakami,K.,
Ishii,S., Kawai,Y., Saito,K., Yamamoto,J., Wakamatsu,A.,
Nakamura,Y., Nagahara,K., Masuko,Y., Ninomiya,K. and Iwayanagi,T.
NEBO human cDNA sequencing project
Unpublished
2 (bases 1 to 2160)
Isegai,T. and Otsuki,T.
Direct Submission
Submitted (23-AUG-2000) Takao Isegai, Helix Research Institute,
Genomics Laboratory, 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
(E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
NEBO human cDNA sequencing project supported by Ministry of
International Trade and Industry of Japan; cDNA full insert
sequencing: Research Association for Biotechnology; cDNA library
construction: 5'- & 3'-end one pass sequencing and clone selection:
Helix Research Institute (supported by Japan Key Technology Center
etc.) and Department of Virology, Institute of Medical Science,
University of Tokyo.
FEATURES
source
Location/Qualifiers
1..2160
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MAMMA1001476"
/tissue_type="mammary gland"
/clone_1db="MAMMA1"
/note="cloning vector: PME18SFLJ3"
95..928
/note="unnamed protein product"
/codon_start=1
/protein_id="BAB14010.1"
/db_xref="GI:10433688"
/translation="MASAGBDCSPAEADRPQRPFLIGVSGTASGKSTVCEKIM
ELIGONEVEORQKRVILSDRFYKVLTAEOKAKALGOYFHDPAFDDMLHRTLK
NIVGKTYEVTYDPTVHSRLPETTVYPAIVLPEGLIYVSEIEMPHILRLFVDT
DSYRLSRVLRVDRGRDLQILITQYTPYKPAFEECLPTKRYAIVIIIPRGDMV
AIVNLVQHIQDILMGDICKMHRGSGNSYKRTPEPDHGMULTSGRSHLESSRP
H"
CDS
BASE COUNT 457 a 591 c 671 g 441 t
ORIGIN
Query Match 98.5%; Score 1599.4; DB 9; Length 2160;
Best Local Similarity 99.7%; Freq. No. 0;
Matches 1614; Conservative 0; Mismatches 1; Indels 4; Gaps 1;
QY 6 GTCCGCTCGACCTGCGGCTGGGGCGGCGCGGCGCGGGAAGGGCGGCGGCGG 65
DB |||||
QY 7 GTCCGCTCGACCTGCGGCTGGGGCGGCGCGGCGCGGGAAGGGCGGCGGCGG 66
DB |||||
QY 66 GACCCGATGCGCGGAGCGGAGCGGAGATGCTTCGCGGAGGAGCGGAGCTCGAGAG 125
DB |||||
QY 67 GACCCGATGCGCGGAGCGGAGCGGAGATGCTTCGCGGAGGAGCGGAGCTCGAGAG 126
DB |||||
QY 126 CCCCCGCGGAGCGGAGCGGAGCGGAGATGCTTCGATGAGGAGTGAAGCGGCGG 185
DB |||||
QY 127 CCCCCGCGGAGCGGAGCGGAGCGGAGATGCTTCGATGAGGAGTGAAGCGGCGG 186
DB |||||
QY 186 CACTGCAGCGGAGGAGTGCACCGGTGTGTGAGGAATCATGTAGTTGCTGGAGAGAA 245
CGA 245

DB |||||
QY 187 CACTGCAGCGGAGGAGTGCACCGGTGTGTGAGGAATCATGTAGTTGCTGGAGAGAGCA 246
CGA 246
QY 246 GTTGAACAGCGGAGCGGAGGAGTGTCTATCTGAGCCAGAGAGTTCTACAAGTCT 305
DB |||||
QY 247 GTTGAACAGCGGAGCGGAGGAGTGTCTATCTGAGCCAGGAGTGTCTACAAGTCT 306
DB |||||
QY 306 GACGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 365
DB |||||
QY 307 GACGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 366
DB |||||
QY 366 CTTTGAATGATTTGATGACAGAGCTGTGAAGAACATCTGTGAGAGGAGGAGGAGGAG 425
DB |||||
QY 367 CTTTGAATGATTTGATGACAGAGCTGTGAAGAACATCTGTGAGAGGAGGAGGAGGAG 426
DB |||||
QY 426 GTTGAACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 485
DB |||||
QY 427 GTTGAACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 486
DB |||||
QY 486 TGGGAGCGTGTCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 545
DB |||||
QY 487 TGGGAGCGTGTCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 546
DB |||||
QY 546 GTTCACTGCGGCTCTGCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 605
DB |||||
QY 547 GTTCACTGCGGCTCTGCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 606
DB |||||
QY 606 CCGGAGCGTGTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 665
DB |||||
QY 607 CCGGAGCGTGTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 666
DB |||||
QY 666 GAAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 725
DB |||||
QY 667 GAAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 726
DB |||||
QY 726 ACAGAGAGTGAACATATGTTGTCATCACTGATCTGACAGACATCAAGAGCATTTCT 785
DB |||||
QY 727 ACAGAGAGTGAACATATGTTGTCATCACTGATCTGACAGACATCAAGAGCATTTCT 786
DB |||||
QY 786 GAATGATGATCTGCAATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 845
DB |||||
QY 787 GAATGATGATCTGCAATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 846
DB |||||
QY 846 CTTTTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 905
DB |||||
QY 847 CTTTTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 906
DB |||||
QY 906 GTCCAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 965
DB |||||
QY 907 GTCCAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 966
DB |||||
QY 966 TGTGTGTTACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1025
DB |||||
QY 967 TGTGTGTTACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1026
DB |||||
QY 1026 ACCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1085
DB |||||
QY 1027 ACCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1086
DB |||||
QY 1086 CACTCAGACTCACTTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1145
DB |||||
QY 1087 CACTCAGACTCACTTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1146
DB |||||
QY 1146 CTTGTTGCGGTTAAAGATCTCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1205
DB |||||
QY 1147 CTTGTTGCGGTTAAAGATCTCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1206
DB |||||
QY 1206 CTTGTTGCGGTTAAAGATCTCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1265
DB |||||
QY 1207 CTTGTTGCGGTTAAAGATCTCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1266
DB |||||
QY 1266 AGACACTGCTGCTGATGATTTTGTGACAGTGAACCACTGTGAGAGAGATTTTTC 1325
GAT 1325

Db	1267	AGACACTGCGCGGCTGATGTTTGTGTGACAGTGAACCCACAGTGGGAGAGAGTTTTC	1328b
Qy	1326	CAGTCGTAGTCGTGTTCTTACACACTGACACACATTACTCAAAAGTTTGTGAACAAGTAC	1385b
Db	1327	CAGTCGTAGTCGTGTTCTTACA-----CACACACATTAATCAAAAGTTTGTGAACAAGTAC	1382b
Qy	1386	TTTCCCTTTTACATNGTTACATGTGCTCATGTGTTTCTGTGTTTCTGTTCATTAACAACAGG	1445b
Db	1383	TTTCCCTTTTACACATGTTACATGTGCTCATGTGTTTCTGTGTTTCTGTTCATTAACAACAAG	1442b
Qy	1446	CTGGTTGTGGCCTTCAAACTTAATTTACATGAACCCAGTGGTTGGAGTCACAGTGGCCTTA	1505b
Db	1443	CTGGTTGTGGCCTTCAAACTTAATTTACATGAACCCAGTGGTTTGCAGTCCAGCGTGCCCTTA	1502b
Qy	1506	CACGGATATGCGGAGCCACTGAGGGATGTTTCCGCCCTTGCTGTGCTTTAAAGCCAGA	1565b
Db	1503	CACGGATATGCGGAGCCACTGAGGGATGTTTCCGCCCTTGCTGTGCTTTAAAGCCAGA	1562b
Qy	1566	GAAGCGAGGCGGATGCTCTGGAGAGCCACGATATACCCAGGCTTTGGGGGGCCAG	1624b
Db	1563	GAAGCGAGGCGGATGCTCTGGAGAGCCACGATATACCCAGGCTTTGGGGGGCCAG	1621b
RESULT 4			
AK057848			
LOCUS	AK057848	2228 bp	mRNA
DEFINITION	Homo sapiens CDNA FLJ25119 fl1, clone CBR05878, highly similar to		
ACCESSION	AK057848		
VERSION	AK057848.1	GI:16553809	
KEYWORDS	oligo capping; fls (full insert sequence).		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 Tashiro, H., Yamazaki, M., Watanabe, K., Kumagai, A., Itakura, S., Fukuzumi, Y., Fujimori, Y., Komiyama, M., Suzuki, Y., Hata, H., Nakagawa, K., Morigano, S., Morigano, M., Kawamura, M., Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Nishikawa, T., Sugiyama, A., Kawakami, B., Nagai, K., Isogai, T. and Sugano, S.		
TITLE	NEDO human cDNA sequencing project		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 2228)		
AUTHORS	Sugano, S. and Suzuki, Y.		
TITLE	Direct Submision		
JOURNAL	Submitted (24-OCT-2001) Sumio Sugano, Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure, Human Genome Center; Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan (E-mail: cdna@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286, Fax:81-3-5449-5416)		
COMMENT	NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan. cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction and 5'-end and one pass sequencing: Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure, Human Genome Center; 3'-end one pass sequencing: RAB; clone selection from full insert sequencing: RAB and Helix Research Institute.		
FEATURES	Location/Qualifiers		
SOURCE	1..2228		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="CBR05878"		
	/tissue_type="brain"		
	/clone_id="CBR"		
	/note="Cloning vector: pME18SFL3"		
BASE COUNT	478 a 598 c 696 g 456 t		
ORIGIN			
Query Match	89.5%	Score 1453.4	DB 9; Length 2228;
Best Local Similarity	93.7%	Pred. NO. 1.9e-296;	
Matches 1589; Conservative	0; Mismatches 1;	Indels 105; Gaps 3;	

QY	31	GGGCGCCGCGGCGCGGGAGAGGGGCGCGCGGGGACCCGATGCGCGGAGCGGAGGCC	90
Db	1	GGGGCGCCCGGGCCTGTGGGGAAGGGGCGGGCGCGGGGAGCCGATGCGCGGGAGCGAGGCC	60
QY	91	GAGATGGCTTCGCGGGGAGCGAGAGCTGCGAGAGCCCGCGCGCGGAGGCCGATCCG	150
Db	61	GAGATGGCTTCGCGGGGAGCGAGAGCTGCGAGAGCCCGCGCGCGGAGGCCGATCCG	120
QY	151	CACCAAGCGCCCTTCTGTATAGGGGTAGCGGCGGCACTGCGAGCGGAG-----	201
Db	121	CACCAAGCGCCCTTCTGTATAGGGGTAGCGGCGGCACTGCGAGCGGAGAACCGAGCG	180
QY	202	-----TCGACCGGTGTGGAAGAAATCATGGAG	228
Db	181	TTCCAGGCGGGGAAATCCGCGCTTGTGTCTCCAGTCCGCGTGTGAGAAATCATGGAG	240
QY	229	TTGCTGGGAGCAGAACGAGGTGGAACAGCGGAGCGGAGGTGTATCTTGAGCGAGAGC	288
Db	241	TTGCTGGGAGCAGAACGAGGTGGAACAGCGGAGCGGAGGTGTATCTTGAGCGAGAGC	300
QY	289	AGGTTCTAACAGGTCTCTGACGCGAGAGCAGAGGCCCAAGCCTTGAAAGCACTACAT	348
Db	301	AGGTTCTAACAGGTCTCTGACGCGAGAGCAGAGGCCCAAGGCTTGAAAGCACTACAT	360
QY	349	TTTGAACATCCAGATGCGCTTTGATATATGATTGTATGACAGAGCTCTGAAGAAATCGTG	408
Db	361	TTTGAACATCCAGATGCGCTTTGATATATGATTGTATGACAGAGCTCTGAAGAAATCGTG	420
QY	409	GAGGCGAAAACGATGAGAGGTGCGACCTATGATTTTGTACACACTCAAGTTACAGAG	468
Db	421	GAGGCGAAAACGATGAGAGGTGCGACCTATGATTTTGTACACACTCAAGTTACAGAG	480
QY	469	ACCAAGGTGTCTACCCGCGGGAGCGAGGTCTGTATAGGGGCACTTGAGTTCTACAGC	528
Db	481	ACCAAGGTGTCTACCCGCGGGAGCGAGGTCTGTATAGGGGCACTTGAGTTCTACAGC	540
QY	529	CAGAGATTCGCGGACATGTTTCACTCTGCGCTCTCTGTGACACACGACTCCGAGCTCAGG	588
Db	541	CAGAGATTCGCGGACATGTTTCACTCTGCGCTCTCTGTGACACACGACTCCGAGCTCAGG	600
QY	589	CTGTCTCGAAGAGTTCTTCGCGGAGCGTGCGCTCGAGGGAGGACCTGAGAGCAATTTCTAGC	648
Db	601	CTGTCTCGAAGAGTTCTTCGCGGAGCGTGCGCGCGAGGGAGGACCTGAGAGCAATTTCTAGC	660
QY	649	CAGTACACCACTTCGTGGAAGCGGCGCTTCGAGGAGTTCTGCGCTGCGG-----	696
Db	661	CAGTACACCACTTCGTGGAAGCGGCGCTTCGAGGAGTTCTGCGCTGCGCGAGAGCATCTGA	720
QY	697	-----ACAAAGAGTATG	709
Db	721	CAGGGAATGAGAGTCAGCATTTGAGCCATGATGTTGTGATGAGGGAACAAAGAGATATG	780
QY	710	CCGATGTGATCATCCCAAGAGAGATGGAACAATAGTTGTGCATCACTGATCGTGGAGC	769
Db	781	CCGATGTGATCATCCCAAGAGAGATGGAACAATAGTTGTGCATCACTGATCGTGGAGC	840
QY	770	ACATCCAGGACATTTCTGATGTGTGACATCTTGCAATGGAACCGAGGAGGTCCATGGGC	829
Db	841	ACATCCAGGACATTTCTGATGTGTGACATCTTGCAATGGAACCGAGGAGGTCCATGGGC	900
QY	830	GGAGCTACAGCGGACCTTTTCTGAGCCAGGGGACCAACCTTGAGGATGCTGACCTCTGGCA	889
Db	901	GGAGCTACAGCGGACCTTTTCTGAGCCAGGGGACCAACCTTGAGGATGCTGACCTCTGGCA	960
QY	890	AACGTCACATTTTGAATGTCAGACGACAGACCCCACTAGGGGGCTGCGAGCCTCAGGGCA	949
Db	961	AACGTCACATTTTGAATGTCAGACGACAGACCCCACTAGGGGGCTGCGAGCCTCAGGGCA	1020
QY	950	GGTCTCCCGCCCGGCAATGTGTTCAGGAGCTGAGCTGGGAGATGCCACCCACATCCAC	1009
Db	1021	GGTCTCCCGCCCGGCAATGTGTTCAGGAGCTGAGCTGGGAGATGCCACCCACATCCAC	1080

QY 1010 TCGTCTCTGCGGCGCACCCAGAGGAGTGTAGCAGCGAGCCCTTCTCACTCAGAGT 1069
 DB 1081 TCGTCTCTGCGGCGCACCCAGAGGAGTGTAGCAGCGAGCCCTTCTCACTCAGAGT 1140
 QY 1070 GGAACCTCAGATGTGTCTACTGACTCAACTTGTCTGGACACTGACGAGCGTCTCTGAGG 1129
 DB 1141 GGAACCTCAGATGTGTCTACTGACTCAACTTGTCTGGACACTGACGAGCGTCTCTGAGG 1200
 QY 1130 TTTTCAGCCACTTAGGCTCGTTGCGGTTTAAAGATCCCTTAGGCACTGAGAAATGCA 1189
 DB 1201 TTTTCAGCCACTTAGGCTCGTTGCGGTTTAAAGATCCCTTAGGCACTGAGAAATGCA 1260
 QY 1190 CAGAAATGTCAGAGAACCTGAGAGCTTCTGTAGAGAAATGAGGACATTAATTGGGAA 1249
 DB 1261 CAGAAATGTCAGAGAACCTGAGAGCTTCTGTAGAGAAATGAGGACATTAATTGGGAA 1320
 QY 1250 ATTGAGGAGACAGCTTAACACTGCTGCTGCTGATGTTTGTGTGACATGAAACCAAGT 1309
 DB 1321 ATTGAGGAGACAGCTTAACACTGCTGCTGCTGATGTTTGTGTGACATGAAACCAAGT 1380
 QY 1310 GGGAGAGAGTTTTCAGTCTGATCTGCTTACACACTGACACATTAATCAAAAG 1369
 DB 1381 GGGAGAGAGTTTTCAGTCTGATCTGCTTACACACTGACACATTAATCAAAAG 1436
 QY 1370 TTTTGTGAACAAGTACTTCTTTTTCATGTTACATGCTCTCATGTTTCTGTTTCT 1429
 DB 1437 TTTTGTGAACAAGTACTTCTTTTTCATGTTACATGCTCTCATGTTTCTGTTTCT 1496
 QY 1430 GTTTCATTAACACAGGCTGTGTGTGCTTACAACTTAATTCATGACCCAGTGTGTTGC 1489
 DB 1497 GTTTCATTAACACAGGCTGTGTGTGCTTACAACTTAATTCATGACCCAGTGTGTTGC 1556
 QY 1490 AGTCCAGGCTGAGCTTACACGATATGGGAGCACTGAGGATGTTTCCCCCTGTTGCTT 1549
 DB 1557 AGTCCAGGCTGAGCTTACACGATATGGGAGCACTGAGGATGTTTCCCCCTGTTGCTT 1616
 QY 1550 GTGCTTTAAGGACAGAGCGAGCGGATGCTCTGAAAGCAACCCAGCATCAACCCAGG 1609
 DB 1617 GTGCTTTAAGGACAGAGCGAGCGGATGCTCTGAAAGCAACCCAGCATCAACCCAGG 1676
 QY 1610 CTTGTGCGGGGCGCAG 1624
 DB 1677 CTTGTGCGGGGCGCAG 1691
 RESULT 5
 LOCUS BC015547 2072 bp mRNA linear PRI 04-OCT-2001
 DEFINITION Homo sapiens, similar to uridine-cytidine kinase 1, clone MGC.9668
 IMAGE:3845821, mRNA, complete cds.
 ACCESSION BC015547
 VERSION BC015547.1 GI:15930229
 KEYWORDS MGC.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 2072)
 AUTHORS Strausberg, R.
 TITLE Direct Submission
 JOURNAL Submitted (01-OCT-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
 COMMENT Contact: MGC help desk
 Email: cgapbs-rt@mail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome
 Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: http://www-shgc.stanford.edu

Contact: (Dickson, Mark) mcd@paxil.stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
 R. M.
 Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
 Series: IRAX Plate: 20 Row: 5 Column: 5
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 13899252.
 FEATURES
 source
 1..2072
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:9668 IMAGE:3845821"
 /tissue_type="Colon, adenocarcinoma"
 /clone_id="NIH_MGC_65"
 /lab_host="DH10B"
 /note="vector: pCMV-Sport6"
 87..692
 /codon_start=1
 /product="Similar to uridine-cytidine kinase 1"
 /protein_id="AAH15547.1"
 /db_xref="GI:15930230"
 /translation="MASAGEDESPAPREAPRPHQPLIGVSGTASGKTVCEKIM
 ELIGONEVEORRKYVILSDRFYKVLAEOKKALKGOYFNPDPAPFNDLHRTLK
 NIVEGKVEPVDVPLFTHSLPETTVVYPAVDVLFGSLIVFYSQELRDMFHLRLPVDL
 DSPVRSRBRKEVRCRDHPFRSQYQCHQPDRAAHGHSEW"
 BASE COUNT 455 a 559 c 637 g 421 t
 ORIGIN
 Query Match 86.0%; Score 1397.4; DB 9; Length 2072;
 Best Local Similarity 93.8%; Pred. No. 1,3e-284;
 Matches 1517; Conservative 0; Mismatches 1; Indels 99; Gaps 2;
 QY 8 CGCCTCCGACCTTGGCGCTGGCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGGGA 67
 DB 1 CGCCTCCGACCTTGGCGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGGGA 60
 QY 68 CCCGATGCGGGGAGCGGAGCGGAGATGCGCTTCCGCGGAGCGGAGCGGAGCGGAGCGC 127
 DB 61 CCCGATGCGGGGAGCGGAGCGGAGATGCGCTTCCGCGGAGCGGAGCGGAGCGGAGCGC 120
 QY 128 CCGGCGCGGAGCGGAGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGGCA 187
 DB 121 CCGGCGCGGAGCGGAGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGGCA 180
 QY 188 CTGCGACGCGGAGAGTCGACCGTGTGTGAGAAATCATGAGTTCTGGGACAGACGAGG 247
 DB 181 CTGCGACGCGGAGAGTCGACCGTGTGTGAGAAATCATGAGTTCTGGGACAGACGAGG 240
 QY 248 TGGAAACGCGGAGCGGAGCGGAGTGTGATCCTGAGCGACAGGACAGTTCTCAAGTCTGA 307
 DB 241 TGGAAACGCGGAGCGGAGCGGAGTGTGATCCTGAGCGACAGGACAGTTCTCAAGTCTGA 300
 QY 308 CGGACGAGCGAGAGCGCAAGCGCTTGAAGAAGATGATCAATTTGACCATCCAGATGCTT 367
 DB 301 CGGACGAGCGAGAGCGCAAGCGCTTGAAGAAGATGATCAATTTGACCATCCAGATGCTT 360
 QY 368 TTGATAATGATTTGATGACAGAGCTTGAAGAATCGTGAAGGCGCAAAACGCTGAGG 427
 DB 361 TTGATAATGATTTGATGACAGAGCTTGAAGAATCGTGAAGGCGCAAAACGCTGAGG 420
 QY 428 TGGCGACCTATGATTTTGTGACACACTCAAGGTTTACCAAGACCAAGCGTGTCTACCTG 487
 DB 421 TGGCGACCTATGATTTTGTGACACACTCAAGGTTTACCAAGACCAAGCGTGTCTACCTG 480
 QY 488 CGGACGAGTCTTCTGTTTGAAGGACATCTGAGTGTCTTACAGCGAGGATCCGGAACATGT 547
 DB 481 CGGACGAGTCTTCTGTTTGAAGGACATCTGAGTGTCTTACAGCGAGGATCCGGAACATGT 540
 QY 548 TCGACCTGCGCTCTTCTGTGACACCGACCTCGAGCTGAGCTGTCTCGAAGATCTCTCC 607

```

Db      541 TCCACCTGCGCCTCTCTCTGTGACACCGATCCGAGCTGCTGTCGAAGA----- 593
Qy      608 GGGAGCTGCGCGGAGGGAGGAGCCTTGAGCAGATTCTACCGCAGTACACACTTTCGTGA 667
Db      594 ----- 593
Qy      668 AGCCGCGCTTCGAGGAGTTCTGCTGCCGACAAAGAAATATGCGATGTGATCCAC 727
Db      594 -----GACAAAGAAATATGCCGATGTGATCCAC 625
Qy      728 GAGAGTGTGACAATGTGTTGCCATCAACTGATGTGTGACAGACATCCAGACATTCTGA 787
Db      626 GAGAGTGTGACAATGTGTTGCCATCAACTGATGTGTGACAGACATCCAGACATTCTGA 685
Qy      788 ATGGTGCATCTGCAAAATGTGACCGAGGAGGTTCCAAATGGGCGGAGCTACAAAGGAGACT 847
Db      686 ATGGTGCATCTGCAAAATGTGACCGAGGAGGTTCCAAATGGGCGGAGCTACAAAGGAGACT 745
Qy      848 TTTCTGAGCCAGGGGAGCACCCTGGGATGTGACCTCTGCGAAAACGTCACATTGGAGT 907
Db      746 TTTCTGAGCCAGGGGAGCACCCTGGGATGTGACCTCTGCGAAAACGTCACATTGGAGT 805
Qy      908 CCAGCAGACAGACCCCACTGAGGGGCTGCGGAGCTCAAGGGGAGGTTCTCCGCGCGCATG 967
Db      806 CCAGCAGACAGACCCCACTGAGGGGCTGCGGAGCTCAAGGGGAGGTTCTCCGCGCGCATG 865
Qy      968 TGTGTTCAGGGAGCTGAGCGCTGGGGAGCGCCCAACCCCACTGCTTCTCTCGGCGCAC 1027
Db      866 TGTGTTCAGGGAGCTGAGCGCTGGGGAGCGCCCAACCCCACTGCTTCTCTCGGCGCAC 925
Qy      1028 CCCAGGGGAGTGTGAGCAGGAGCCTTCTCTCACTGAGGAGTGGAAAACAGATGTGTCA 1087
Db      926 CCCAGGGGAGTGTGAGCAGGAGCCTTCTCTCACTGAGGAGTGGAAAACAGATGTGTCA 985
Qy      1088 CTCGACTCAACTCTGCTGGGACACTGACAGGCGTTTCTGAGGTTTTCAGCCACTTAGGCT 1147
Db      986 CTCGACTCAACTCTGCTGGGACACTGACAGGCGTTTCTGAGGTTTTCAGCCACTTAGGCT 1045
Qy      1148 CGTTGGCGGTTTAAAGATCCCTCTAGTCACTGAAATATGCCAAGATGTGACAGAAAGCC 1207
Db      1046 CGTTGGCGGTTTAAAGATCCCTCTAGTCACTGAAATATGCCAAGATGTGACAGAAAGCC 1105
Qy      1208 TGGGAGGCTTCTGTGAGGAAATGTGAGGACATTAATTGGGGGAAATTGAGAGACAGCCTAG 1267
Db      1106 TGGGAGGCTTCTGTGAGGAAATGTGAGGACATTAATTGGGGGAAATTGAGAGACAGCCTAG 1165
Qy      1268 ACACCTGCTGGCTGATGTTTGTGAACGTGAACCAACAGTGGAGAGAGTTCCTCA 1327
Db      1166 ACACCTGCTGGCTGATGTTTGTGAACGTGAACCAACAGTGGAGAGAGTTCCTCA 1225
Qy      1328 GTCTGATCTGGTCTTACACACTCACACATTACTCAAAAGTTTGTGAAACAAGTACTT 1387
Db      1226 GTCTGATCTGGTCTTACACACTCACACATTACTCAAAAGTTTGTGAAACAAGTACTT 1281
Qy      1388 TCCCTTTTTCATGTTACATGCTCATGTTTCTGTTTCTGTTTCAATAACAAGGCT 1447
Db      1282 TCCCTTTTTCATGTTACATGCTCATGTTTCTGTTTCTGTTTCAATAACAAGGCT 1341
Qy      1448 GGTGTGGCTTAACAACCTAATTTCAATGACCCAGTGGTTTTCAGTCCAGCGTGGCTTACA 1507
Db      1342 GGTGTGGCTTAACAACCTAATTTCAATGACCCAGTGGTTTTCAGTCCAGCGTGGCTTACA 1401
Qy      1508 CGGATATGGGGAAGCACTGAGGGAATGTTTCCCGGCTTGTGCTTGAAGGCAAGAGA 1567
Db      1402 CGGATATGGGGAAGCACTGAGGGAATGTTTCCCGGCTTGTGCTTGAAGGCAAGAGA 1461
Qy      1568 AGGAGGCGGATGCGCTGGAAGACCCAGCATCAACCCAGGCTTGTGCGGGGCGAG 1624
Db      1462 AGGAGGCGGATGCGCTGGAAGACCCAGCATCAACCCAGGCTTGTGCGGGGCGAG 1518

```

RESULT 6
AF125106

```

LOCUS      AF125106                1395 bp    mRNA    linear    PRI 07-FEB-2002
DEFINITION Homo sapiens uridine kinase mRNA, complete cds.
ACCESSION  AF125106
VERSION    AF125106.1 GI:18568108
KEYWORDS
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
REFERENCE  1 (bases 1 to 1395)
AUTHORS   Xin,Y.R., Yu,L. and Zhao,S.Y.
TITLE      Cloning of a new human cDNA similar to Mus musculus uridine kinase
           mRNA
JOURNAL    Unpublished
REFERENCE  2 (bases 1 to 1395)
AUTHORS   Ding,J.B., Yu,L. and Zhao,S.Y.
TITLE      Direct Submission
JOURNAL    Submitted (02-FEB-1999) Lab of Human Gene Research, Institute of
           Genetics, Fudan University, No. 220 Handan Rd., Shanghai 200433,
           People's Republic of China
FEATURES
     source
         1..1395
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            8..838
            /codon_start=1
            /product="uridine kinase"
            /protein_id="AA175943.1"
            /db_xref="GI:18568109"
            /translation="MASAGSEKARAGANRPHORPFLIGSGTASGKSTVEKIME
            LLEQNEVEQRHGVIVLSODRFYKUTATGAKAKALGQYRFPDPAFNDLHRTLKN
            IVEGKVEVPTDFVTHSRLEPTVVPADVLFESILVFSQELDMPLRLFVTD
            SDVRLSRVLRVLRGRDEQLILOYTTEKPAFEFCELPKKYAVIILPRGDVNA
            INIYVHIDILINDICKWHRGGSNRTYKTFPSBPDHGMILSGKSHLSSSRPH
            "
BASE COUNT 325 a 352 c 411 g 307 t
ORIGIN
Query Match 71.2%; Score 1156.8; DB 9; Length 1395;
Best Local Similarity 94.1%; Pred. No. 8.6e-234;
Matches 1317; Conservative 0; Mismatches 57; Indels 26; Gaps 10;
Qy      87 GGGCGAGATGTGCTTGGCGGAGGCGGAAGACTCGAGAGCCCGCGGAGCCGACCG 146
Db      1 GGGCGAGATGTGCTTGGCGGAGGCGGAAG---GTGGAGAGCCCGCGGAGCCGAAACCG 57
Qy      147 TCCGACACAGCGGCGCTTCTGATAGGGGTGAGCGGCGGACATGCCAGGCGGAAGTCGAC 206
Db      58 TCCGACACAGCGCTTCTTCTGATAGGGGTGAGCGGCGGACATGCCAGGCGGAAGTCGAC 117
Qy      207 CGTGTGTGAGAAGATCATGAGATGTGCTGGGACAGAAAGAGTGGAAACAGCGGACGCGAA 266
Db      118 CGTGTGTGAGAAGATCATGAGATGTGCTGGGACAGAAAGAGTGGAAACAGCGGACGCGTAA 177
Qy      267 GGTGTGATCTCTGAGCGCAGGAGTCTTCAAGAGTCTCTGACGCGGAGAGCAAGGCCAA 326
Db      178 GGTGTGATCTCTGAGCGCAGGAGTCTTCAAGAGTCTCTGACGCGGAGAGCAAGGCCAA 237
Qy      327 GGCCTTGAAGAGCAGTACAAATTTGACCATTCAGATGCGCTTTGATTAATGATTGATGCA 386
Db      238 GGCCTTGAAGAGCAGTACAAATTTGACCATTCAGATGCGCTTTGATTAATGATTGATGCA 297
Qy      387 CAGGACTCTGAAGAACATCGTGGAGGGCAAAAACGATGAGGTGCCAGACTATGATTTGT 446
Db      298 CAGGACTCTGAAGAACATCGTGGAGGGCAAAAACGATGAGGTGCCAGACTATGATTTGT 357
Qy      447 GACACACTGAAGTTTACGAGAGCACACGCTGTACCTCTGGGAGAGTGTCTGTTTGA 506
Db      358 GACACACTGAAGTTTACGAGAGCACACGCTGTACCTCTGGGAGAGTGTCTGTTTGA 417
Qy      507 GGGCATCTGTTGTTCTACGCGAGGAGATCCGGGACATGTTCCACCTGCGCTTCTGCT 566

```

```

Db 418 GGGCATCTTGGTCTTACAGCCAGAGATCCGGAGACATGTTCCACTGCGCTCTTGT 477
Qy 567 GGACCCGACCTCCGACGCTCCTCTCGAAGATTTCTCCGGGACCTGCGCCGAGGAG 626
Db 478 GGACCCGACCTCCGACGCTCCTCTCGAAGATTTCTCCGGGACCTGCGCCGAGGAG 537
Qy 627 GGACCTGGAGAGATTTCTGACGACAGTACACCACTTCTGGAAGCCGCTTCTGAGAGAT 686
Db 538 GGACCTGGAGAGATTTCTGACGACAGTACACCACTTCTGGAAGCCGCTTCTGAGAGAT 597
Qy 687 CTGCTGCGCCGACAAAGATATGCCGATGTATATCCACGAGAGTGAACAATATGAT 746
Db 598 CTGCTGCGCCGACAAAGATATGCCGATGTATATCCACGAGAGTGAACAATATGAT 657
Qy 747 TGCCATCAACCTGATCGTGAAGACATCATCGAGACATTCGATGGTGAATCTGGAATG 806
Db 658 TGCCATCAACCTGATCGTGAAGACATCATCGAGACATTCGATGGTGAATCTGGAATG 717
Qy 807 GCACCGAGAGAGGTCCTAATGGCGGAGCTACAGCGGACCTTTCTGAGCCAGGAGACA 866
Db 718 GCACCGAGAGAGGTCCTAATGGCGGAGCTAATGAGCGGACCTTTCTGAGCCAGGAGACA 777
Qy 867 CCTTGGAGTGTGACCTCTGCGAAACGGTCACTTTGAGTTCAGCAGCAGACCCACTG 926
Db 778 CCTTGGAGTGTGACCTCTGCGAAACGGTCACTTTGAGTTCAGCAGCAGACCCACTG 837
Qy 927 AGGGGCTGCGGACCTCAGGGGAGGTCCTCGGCGGAGTGTGATGAGGACCTGAGCC 986
Db 838 AGGGGCTGCGGACCTCAGGGGAGGTCCTCGGCGGAGTGTGATGAGGACCTGAGCC 897
Qy 987 TGGGACCGCCACCCACACCCACTGCTTCTCTCGGCGGACCCGAGGAGTGTATGAG 1046
Db 898 TGGGACCGCCACCCACACCCACTGCTTCTCTCGGCGGACCCGAGGAGTGTATGAG 957
Qy 1047 CGAGGCTTCTCTCACTCGAGAGTGAACCTAGATGTGTCACTCACTCAACTTGTCTGG 1106
Db 958 CGAGGCTTCTCTCACTCGAGAGTGAACCTAGATGTGTCACTCACTCAACTTGTCTGG 1016
Qy 1107 GACACTGACAGGCGCTTCTGAGGTTTTCAGGCACTTGAAGCTGCTGCGGTTTAAAGATCC 1166
Db 1017 GACACTGACAGGCGCTTCTGAGGTTTTCAGGCACTTGAAGCTGCTGCGGTTTAAAGATCC 1076
Qy 1167 CTCTAGAGTCACTGAGAAATGCAAGAGATGTTGCAAGAAAGCTGCGGAAAGCTTCTGG 1218
Db 1077 CTCTAGAGTCACTGAGAAATGCAAGAGATGTTGCAAGAAAGCTGCGGAAAGCTTCTGG 1136
Qy 1219 TGTGAGAGATGTGAGGCACTTATTTGGGAAA-----TTGAGAGACAGCCTTACACTG 1273
Db 1137 TGAAGGAAATGTGAAGGACATTAATTTGGGCAAAATTGAGAAAGACAGCCTTACACTG 1196
Qy 1274 G---CTGGCTGATGTTTGTGTTGACA-GTGAACCCAGAGTGGG-AGAGAGTTTTCAG 1328
Db 1197 GCGCTGGGCTTGAATTTTGTGTTGACATGTGAACCCAGAGTGAAGAGAGATTTTTCAG 1256
Qy 1329 TCTGATCTGTTTCTTACACACTCAACAACATTAACCAAAAGTTTGTGAACAAGTACTT 1388
Db 1257 TCTGATCTGTTTCTTACA-----CACACACATTAACCAAAAGTTTGTGAACAAGTACTT 1312
Qy 1389 CCTTTTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1448
Db 1313 CCTTTTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1372
Qy 1449 GTTGTGGCTTACAAACCTAA 1468
Db 1373 GTTGTGGCTTACATTA 1392

```

```

RESULT 7
AF237290 1022 bp mRNA linear PRI 02-AUG-2001
LOCUS Homo sapiens uridine-cytidine kinase 1 (UCK1) mRNA, complete cds.
DEFINITION AF237290
ACCESSION AF237290.1 GI:13506764
VERSION

```

```

KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1022)
AUTHORS Van Rompay,A.R., Norda,A., Linden,K., Johansson,M. and Karlsson,A.
TITLE Phosphorylation of uridine and cytidine nucleoside analogs by two
human uridine-cytidine kinases
JOURNAL Mol. Pharmacol. 59 (5), 1181-1186 (2001)
MEDLINE 21203813
PUBMED 11306702
REFERENCE 2 (bases 1 to 1022)
AUTHORS Van Rompay,A.R., Linden,K., Norda,A., Zhu,C., Zheng,X.,
Johansson,M. and Karlsson,A.
TITLE Human uridine-cytidine kinase 1 and 2 : rate limiting enzymes
unpublished
JOURNAL 3 (bases 1 to 1022)
AUTHORS Van Rompay,A.R., Linden,K., Norda,A., Zhu,C., Zheng,X.,
Johansson,M. and Karlsson,A.
TITLE Submitted (19-FEB-2000) IMPI, Clinical Virology, Huddinge
University Hospital, Stockholm 14186, Sweden
FEATURES
source
1. .1022
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1. .1022
/feature="UCK1"
5. .838
/feature="UCK1"
/function="phosphorylates uridine and cytidine to uridine
monophosphate and cytidine monophosphate"
/codon_start=1
/product="uridine-cytidine kinase 1"
/protein_id="P4K28324.1"
/db_xref="GI:13506765"
/translacion="MASAGDESPAPRAPHORPFLIGSGGTSKSTCEKIM
ELIGONBEVROKRVITLSODRFYKULTBQKAKLKGONPBPAPNDLHRLK
NIVEGTVPEPTVDFVTHSLRPETTVVYPADVLFEGILVFSQELRDMFHLFPVDT
DSVRLSRVLRVRRGRDLEQLITQYTFVKAPEFELPTKRYADVIIPREVDMV
AIIILYQHIQDILINDICIKWHRSGNSGRSKYTFSPBGPMLTSGKSHLESSRP
H"
BASE COUNT 227 a 282 c 324 g 189 t
ORIGIN
Query Match 62.7%; Score 1018.8; DB 9; Length 1022;
Best Local Similarity 99.8%; Pred. No. 1.2e-204;
Matches 1020; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

Db      301 GACTGTGAAACATCGTGAAGGGCAAAACGGTGAAGTCCGACCTATGATTTGTGAC 360
Qy      450 ACCTCAAGTTTCCAGAGACCAAGGTGTCTACCCCTGGACCTGTTCTGTTGAGG 509
Db      361 ACATCAAGGTATACAGAGACCAAGGTGTCTACCCCTGGACCTGTTCTGTTGAGG 420
Qy      510 CATTTGGTGTCTACAGAGACCAAGGTGTCTACCCCTGGACCTGTTCTGTTGAGG 569
Db      421 CATTTGGTGTCTACAGAGACCAAGGTGTCTACCCCTGGACCTGTTCTGTTGAGG 480
Qy      570 CACCACTCCGACGTGACGTGTCTGAGAGATTCTCCGGACGTGCGCGAGAGAGGA 629
Db      481 CACCACTCCGACGTGACGTGTCTGAGAGATTCTCCGGACGTGCGCGAGAGAGGA 540
Qy      630 CTTGAGACATTTCTGACAGATCACCACTTTGTGTAAGCCGGCTTCGAGAGTTCTG 689
Db      541 CTTGAGACATTTCTGACAGATCACCACTTTGTGTAAGCCGGCTTCGAGAGTTCTG 600
Qy      690 CTTGAGACATTTCTGACAGATCACCACTTTGTGTAAGCCGGCTTCGAGAGTTCTG 749
Db      601 CTTGAGACATTTCTGACAGATCACCACTTTGTGTAAGCCGGCTTCGAGAGTTCTG 660
Qy      750 CATCAACCTGATGCTGACAGATCACCACTTTGTGTAAGCCGGCTTCGAGAGTTCTG 809
Db      661 CATCAACCTGATGCTGACAGATCACCACTTTGTGTAAGCCGGCTTCGAGAGTTCTG 720
Qy      810 CCGAGAGAGGTTCATAGGGGAGAGCTCAAGCGGACCTTTTCTGAGCCAGGGACCA 869
Db      721 CCGAGAGAGGTTCATAGGGGAGAGCTCAAGCGGACCTTTTCTGAGCCAGGGACCA 780
Qy      870 TGGGATCTGACCTCTGAGCAAGGGTGCATATTTGAGTCCAGAGAGAGAGAGAGG 929
Db      781 TGGGATCTGACCTCTGAGCAAGGGTGCATATTTGAGTCCAGAGAGAGAGAGAGG 840
Qy      930 GGGCTGCGAGCCTTCAAGGAGAGGTCTCCCGCCGATGTGTCTGAGAGAGAGAGG 989
Db      841 GGGCTGCGAGCCTTCAAGGAGAGGTCTCCCGCCGATGTGTCTGAGAGAGAGAGG 900
Qy      990 GGAAGGCCACCCACACCACTGCTTCTCTGAGGAGAGAGAGAGAGAGAGAGG 1049
Db      901 GGAAGGCCACCCACACCACTGCTTCTCTGAGGAGAGAGAGAGAGAGAGAGG 960
Qy      1050 GGGCTTCTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGG 1109
Db      961 GGGCTTCTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGG 1020
Qy      1110 AC 1111
Db      1021 AC 1022

RESULT 8
AL358781/c 147492 bp DNA linear PRI 06-OCT-2001
LOCUS      Human DNA sequence from clone RP11-33406 on chromosome 9, complete
DEFINITION
ACCESSION AL358781.19 GI:13751418
VERSION    AL358781.19
KEYWORDS   HTG.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 147492)
AUTHORS    Corby,N.
TITLE      Direct Submission
JOURNAL    Submitted (06-OCT-2001) Sanger Centre, Hinxton, Cambridgeshire,
           CB10 1SA, UK. E-mail enquiries: humgeny@sanger.ac.uk
COMMENT    requests: cloneenquiries@sanger.ac.uk
           On Apr 21, 2001 this sequence version replaced gi:13396549.
           During sequence assembly data is compared over overlapping clones.
           Where differences are found these are annotated as variations
           together with a note of the overlapping clone name. Note that the

```

```

variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em: EMBL; Sw:
SWISSPROT; Tr: TrEMBL; Wp: WORMPEP; Information on the WORMPEP
database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr9
RP11-33406 is from the library RPCT-11.2 constructed by the group
of Pieter de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pBACe3.6
This sequence is the entire insert of clone RP11-33406. The true
left end of clone RP11-40A7 is at 113870 in this sequence. The true
right end of clone RP11-643E14 is at 63282 in this sequence.

FEATURES
     source
       1..147492
         /organism="Homo sapiens"
         /mol_type="genomic DNA"
         /db_xref="taxon:9606"
         /chromosome="9"
         /clone="RP11-33406"
         /clone_1ib="RPCT-11.2"
         /note="Sequence from overlapping clone RP11-643E14
         (AL354855). Assembly confirmed by restriction digest."

BASE COUNT  32992 a 36290 c 37577 g 40633 t

ORIGIN
Query Match      54.2%; Score 880; DB 9; Length 147492;
Best Local Similarity 100.0%; Pred. No. 2.4e-175;
Matches 880; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      745 GTTCCATCAACTGATGCTGACGACATTCAGAGAGAGAGAGAGAGAGAGG 804
Db      129488 GTTCCATCAACTGATGCTGACGACATTCAGAGAGAGAGAGAGAGAGAGG 804
Qy      805 TGGCAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGG 864
Db      129428 TGGCAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGG 864
Qy      865 CACCCTGAGATGCTGACCTCTGAGCAACGATCATTTGAGAGAGAGAGAGAGAGAGG 924
Db      129368 CACCCTGAGATGCTGACCTCTGAGCAACGATCATTTGAGAGAGAGAGAGAGG 924
Qy      925 TGAAGGGGCTGCGAGGCTTCAAGGAGAGAGAGAGAGAGAGAGAGAGAGG 984
Db      129308 TGAAGGGGCTGCGAGGCTTCAAGGAGAGAGAGAGAGAGAGAGAGAGAGG 984
Qy      985 CTTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGG 1044
Db      129248 CTTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGG 1044
Qy      1045 AGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGG 1104
Db      129188 AGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGG 1104
Qy      1105 GGGACATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGG 1164
Db      129128 GGGACATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGG 1164
Qy      1165 CCTCTAGGTCACTGAGAAATGCAAGATGTGACAGAGAGAGAGAGAGAGAGAGG 1224

```



```
Db 129068 CCTCTAGTCACTGAGAAATGCCACAGATGTGAGAAAGCTGGAGGCTTCTGTAG 129009
Qy 1225 GAATGTGAGGCATTTATTGGGAAATTGAGAGACAGCTTAGACATGCTGGCTGTAT 1284
Db 129008 GAAATGTGAGGCATTTATTGGGAAATTGAGAGACAGCTTAGACATGCTGGCTGTAT 128949
Qy 1285 GTTTTGTGACAGTGAACCCACAGTGGAGAGAGTTTTCAGTCTGATCTGGTCTTA 1344
Db 128948 GTTTTGTGACAGTGAACCCACAGTGGAGAGAGTTTTCAGTCTGATCTGGTCTTA 128889
Qy 1345 CACACTCAGACATTAACCTCAAAAGTTTGTGAACATTAATCTTTTAAATGTTA 1404
Db 128888 CACACTCAGACATTAACCTCAAAAGTTTGTGAACATTAATCTTTTAAATGTTA 128829
Qy 1405 CATGTCCTCATGTTTCTGTTTCTGTTTCTATTAACACAGGCTGTTGGCTTAACAAC 1464
Db 128828 CATGTCCTCATGTTTCTGTTTCTGTTTCTATTAACACAGGCTGTTGGCTTAACAAC 128769
Qy 1465 CTAATTTATGATGACCCAGTGTGATGCAATCCAGCGTCTCAACGATATGAGGACAC 1524
Db 128768 CTAATTTATGATGACCCAGTGTGATGCAATCCAGCGTCTCAACGATATGAGGACAC 128709
Qy 1525 TGAGGATGTTTTTCCCTTCTGTTGCTTAAAGCAGAAAGCGAGATGCCCT 1584
Db 128708 TGAGGATGTTTTTCCCTTCTGTTGCTTAAAGCAGAAAGCGAGATGCCCT 128649
Qy 1585 GGAAGCACCCAGCATCAACCCAGGCTTGTGCGGGGCGAG 1624
Db 128648 GGAAGCACCCAGCATCAACCCAGGCTTGTGCGGGGCGAG 128609

RESULT 9
AX449219 834 bp DNA linear PAT 03-JUL-2002
LOCUS AX449219
DEFINITION Sequence 3 from Patent WO0202761.
ACCESSION AX449219
VERSION AX449219.1 GI:21697996
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
REFERENCE
AUTHORS Glucksmann, M. A.
TITLE 57658, a human uridine kinase and uses thereof
JOURNAL Patent: WO 0202761-A 3 10-JAN-2002;
Milleium Pharmaceutical, Inc. (US)
FEATURES
source
1. .834
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

BASE COUNT 194 a 218 c 265 g 157 t
ORIGIN

Query Match 51.4%; Score 834; DB 6; Length 834;
Best Local Similarity 100.0%; Pred. No. 1.3e-165;
Matches 834; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 94 ATGGCTTGGGGGAGGAGGAGAGCTGCGAGAGCCCGCGCGAGCGAGCCGTCGCAC 153
Db 1 ATGGCTTGGGGGAGGAGGAGAGCTGCGAGAGCCCGCGCGAGCGAGCCGTCGCAC 60
Qy 154 CAGCGGCTTCTGATGAGGGGTGAGCGGCGCACTGCGACCGGAAAGTGCACCTGTGT 213
Db 61 CAGCGGCTTCTGATGAGGGGTGAGCGGCGCACTGCGACCGGAAAGTGCACCTGTGT 120
Qy 214 GAGAAGATCATGAGATTCTGGGACAGAAAGAGGTGAACAGCGGCGAGCGGAGAGTGGTC 273
Db 121 GAGAAGATCATGAGATTCTGGGACAGAAAGAGGTGAACAGCGGCGAGCGGAGAGTGGTC 180
Qy 274 ATCTGAGCCAGGAGAGGTTCTACAAAGTCTCTGACGCGAGAGCAGAAAGGCCCTTG 333

Db 181 ATCTGAGCCAGGAGAGGTTCTACAAAGTCTCTGACGCGAGAGCAGAAAGGCCCTTG 240
Qy 334 AAAGACAGTACATTTTGAACCATCCAGATGCTTTGATATGATTTGATGCACAGACT 393
Db 241 AAAGACAGTACATTTTGAACCATCCAGATGCTTTGATATGATTTGATGCACAGACT 300
Qy 394 CTGAGAAACATCGTGAAGGGCAAAAGCGTGAAGTGGCGGACCTATGATTTTGAACAC 453
Db 301 CTGAGAAACATCGTGAAGGGCAAAAGCGTGAAGTGGCGGACCTATGATTTTGAACAC 360
Qy 454 TCAAGTTTACAGAGACCAAGTGTGATCTACCTGCGGACGCTGTTCTGTTGAGGATC 513
Db 361 TCAAGTTTACAGAGACCAAGTGTGATCTACCTGCGGACGCTGTTCTGTTGAGGATC 420
Qy 514 TTGCTGTTTACAGCCAGAGATCCGAGACATGTTCCACTGCGCTCTTCTGAGACAC 573
Db 421 TTGCTGTTTACAGCCAGAGATCCGAGACATGTTCCACTGCGCTCTTCTGAGACAC 480
Qy 574 GACTCGAGGTCAGGCTGTCTGAAAGTCTCCGGAAGTGGCGGAGGAGGACCTG 633
Db 481 GACTCGAGGTCAGGCTGTCTGAAAGTCTCCGGAAGTGGCGGAGGAGGACCTG 540
Qy 634 GAGCAGATTTGACGCGATACACACCTTCTGTAAGCGGCTTTCAGAGAGTTCTGCTG 693
Db 541 GAGCAGATTTGACGCGATACACACCTTCTGTAAGCGGCTTTCAGAGAGTTCTGCTG 600
Qy 694 CCGAACAAGATGATGCCGATGATCATCCACAGAGAGTGAACAATATGTTGCCATC 753
Db 601 CCGAACAAGATGATGCCGATGATCATCCACAGAGAGTGAACAATATGTTGCCATC 660
Qy 754 AACCTGCTGAGCAGCATCCAGGACATTCGTAATGTGACATCTGCAATGGACCGA 813
Db 661 AACCTGCTGAGCAGCATCCAGGACATTCGTAATGTGACATCTGCAATGGACCGA 720
Qy 814 GAGAGTCCAAATGGCGGAGCTCAAGACGACCTTTCTGAGCCAGGAGACACCTGGG 873
Db 721 GAGAGTCCAAATGGCGGAGCTCAAGACGACCTTTCTGAGCCAGGAGACACCTGGG 780
Qy 874 ATGCTGACCTTGTGGCAAAAGGTCACATTTGAGATCCAGAGCAGACCCCACTGA 927
Db 781 ATGCTGACCTTGTGGCAAAAGGTCACATTTGAGATCCAGAGCAGACCCCACTGA 834

RESULT 10
AF254133 834 bp mRNA linear PRI 02-MAY-2001
LOCUS AF254133
DEFINITION Homo sapiens uridine kinase mRNA, complete cds.
ACCESSION AF254133
VERSION AF254133.1 GI:13924749
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
REFERENCE
AUTHORS Ho, Y. S. and Johnson, R. K.
TITLE Human uridine kinase from prostate cancer cell line (LNCap)
JOURNAL Unpublished
2 (bases 1 to 834)
AUTHORS Ho, Y. S. and Johnson, R. K.
TITLE Direct Submission
JOURNAL Submitted (10-APR-2000) Oncology Research, SmithKline Beecham, 709
Swedeland Road, King of Prussia, PA 19406, USA
FEATURES
source
1. .834
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1. .834
/note="protein kinase"
/codon_start=1
/product="uridine kinase"
```


/protein_id="AAK49122.1"
/db_xref="GI:13924750"
/translation="MASAGSEPDSPREADRPHORPFLIGVSGRTASGKTYCEKIM
ELGQNEVEORQRKVVLSQDRFKVLAQKAKALGQNFDPDFDNDLHRTLK
NIVEGKTEVPYTPVTHSLPETTVVPAADVLPBGLVYSQEPIDMHLPLFVDT
DSDVRLSRVLRVRGRDLQILTYTTPKVAEFCELPYKQADAVIIIPRGVDMV
AIVNIYQHIQDILNGDICKWHRGGSNRSYKRTFSERGDHPMLTSKRSHLSSSRP
H"

BASE COUNT 194 a 219 c 265 g 156 t
ORIGIN

Query Match 51.3%; Score 832.4; DB 9; Length 834;
Best Local Similarity 99.9%; Pred. No. 2.9e-165;
Matches 833; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 94 ATGGCTTGGCGGGAGGCGAAGACTGCGAAGACCCCGCGCGAGCGGACCGTCCGAC 153
DB 1 ATGGCTTGGCGGGAGGCGAAGACTGCGAAGACCCCGCGCGAGCGGACCGTCCGAC 60
QY 154 CAGCGGCGCTTCTCGATAGGGGGTGAAGGGCGGCACTGCGACCGGGAAGTGCACCGTGT 213
DB 61 CAGCGGCGCTTCTCGATAGGGGGTGAAGGGCGGCACTGCGACCGGGAAGTGCACCGTGT 120
QY 214 GAGAAATCATGAGATTCTGCGACAGAACGAGGTGGAACAGCGCGACCGGAAGTGTGTC 273
DB 121 GAGAAATCATGAGATTCTGCGACAGAACGAGGTGGAACAGCGCGACCGGAAGTGTGTC 180
QY 274 ATCTTGACCCAGACAGAGTCTTCAAGAGTCTTGAAGAGACAGAAAGCCGCTTG 333
DB 181 ATCTTGACCCAGACAGAGTCTTCAAGAGTCTTGAAGAGACAGAAAGCCGCTTG 240
QY 334 AAAGGAGTACATTTTGAACATCCAGATGCTTGAATATGATTTGATGACACAGACT 393
DB 241 AAAGGAGTACATTTTGAACATCCAGATGCTTGAATATGATTTGATGACACAGACT 300
QY 394 CTGAAAGAACATCGTGAAGGCGAAACCGTGAAGGTGCGACCTATGATTTTGTGACAC 453
DB 301 CTGAAAGAACATCGTGAAGGCGAAACCGTGAAGGTGCGACCTATGATTTTGTGACAC 360
QY 454 TCAAGTTACAGAGACACAGGTGTCTACCTGCGGACGTGTCTGTTGAGGGATC 513
DB 361 TCAAGTTACAGAGACACAGGTGTCTACCTGCGGACGTGTCTGTTGAGGGATC 420
QY 514 TTGGTGTTCACAGCCGAGAGATCCGGGACATGTTCCACTGCGCTCTTCTGTGACACC 573
DB 421 TTGGTGTTCACAGCCGAGAGATCCGGGACATGTTCCACTGCGCTCTTCTGTGACACC 480
QY 574 GACTCCGAGCTCAGGCTGTCTCGAAGAGTTCTCCGGGACGTGCGCGAGGGAAGGACCTG 633
DB 481 GACTCCGAGCTCAGGCTGTCTCGAAGAGTTCTCCGGGACGTGCGCGAGGGAAGGACCTG 540
QY 634 GAGCAGATTCTGACAGACACCACTTCTGTGAAGCGGCGCTTGAAGAGTTCTGCTG 693
DB 541 GAGCAGATTCTGACAGACACCACTTCTGTGAAGCGGCGCTTGAAGAGTTCTGCTG 600
QY 694 CCGACAAAGAGATGCGGATGTGATCTCCACGAGAGTGGACATATGTTGCGATC 753
DB 601 CCGACAAAGAGATGCGGATGTGATCTCCACGAGAGTGGACATATGTTGCGATC 660
QY 754 AACCTGATGTCAGACATTCAGAGACATTCGATGATGTCATCTGCAATGGACCGA 813
DB 661 AACCTGATGTCAGACATTCAGAGACATTCGATGATGTCATCTGCAATGGACCGA 720
QY 814 GAGAGGTTCATGGGCGGAGCTTCAAGCGGACCTTTTCTGAGCGGAGGACCACTTGGG 873
DB 721 GAGAGGTTCATGGGCGGAGCTTCAAGCGGACCTTTTCTGAGCGGAGGACCACTTGGG 780
QY 874 ATGCTGACCTCTGGCAACCGTCAATTTGAGAGTCGACGACGACACCCCACTGA 927
DB 781 ATGCTGACCTCTGGCAACCGTCAATTTGAGAGTCGACGACGACACCCCACTGA 834

RESULT 11

BD146824
LOCUS BD146824 753 bp DNA linear PAT 17-JAN-2003
DEFINITION Primer for synthesizing full-length cDNA and use thereof.
ACCESSION BD146824
VERSION BD146824.1 GI:27852582
KEYWORDS JP 2002191363-A/1667.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Oca,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
TITLE Primer for synthesizing full-length cDNA and use thereof
JOURNAL Patent: JP 2002191363-A 1667 09-JUN-2002;
HELIK RESEARCH INSTITUTE
OS Homo sapiens (human)
PN JP 2002191363-A/1667
PD 09-JUN-2002
PP 28-JUN-2000 JP 2000280990
PI TOSHIO OTR, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU
PI SATTO,
PI JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,
PI KEIICHI NAGAI, TETSUJI OTSUKI
PC C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC
10, C12P21/02, C12Q1/68//C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC
PC Primer for synthesizing full-length cDNA and use thereof FH Key
Location/Qualifiers
FT source 1..753
FT Location/Qualifiers
1..753
/organism="Homo sapiens"
/orgnism="Homo sapiens (human)".

BASE COUNT 158 a 196 c 261 g 135 t 3 others
ORIGIN

Query Match 43.7%; Score 710; DB 6; Length 753;
Best Local Similarity 97.9%; Pred. No. 2.1e-139;
Matches 729; Conservative 0; Mismatches 13; Indels 3; Gaps 1;

QY 6 GTGCTTCGACCTCGGCGTGGGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 65
DB 7 GTGCTTCGACCTCGGCGTGGGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 66
QY 66 GACCGATGCGGCGGAGCGGAGCGGAGTGTGCTGGGCGGAGCGGAGACTGCGAGAG 125
DB 67 GACCGATGCGGCGGAGCGGAGCGGAGTGTGCTGGGCGGAGCGGAGACTGCGAGAG 126
QY 126 CCGCGCGGAGGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 185
DB 127 CCGCGCGGAGGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 186
QY 186 CACTGCCAGGCGGAGTGCACCGTGTGTGAAGATGATGAGTCTCTGGACAGACGA 245
DB 187 CACTGCCAGGCGGAGTGCACCGTGTGTGAAGATGATGAGTCTCTGGACAGACGA 246
QY 246 GGTGGAACGCGGACCGGAGAGGTGTGATCTCGAGCGGAGGAGGTTCTACAAAGTCT 305
DB 247 GGTGGAACGCGGACCGGAGAGGTGTGATCTCGAGCGGAGGAGGTTCTACAAAGTCT 306
QY 306 GACGCGAGGACGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGG 365
DB 307 GACGCGAGGACGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGG 366
QY 366 CTTGATTAATGATTTGATGACAGACTCTGAAAGACATCTGTGAGAGGCGAAGCGGTGA 425
DB 367 CTTGATTAATGATTTGATGACAGACTCTGAAAGACATCTGTGAGAGGCGAAGCGGTGA 426
QY 426 GGTGCGGACCTATGATTTTGTGACACACTCAAGGTTACAGAGACCAAGGTGTCTACCC 485

```

Db      427  GGTGCCGACCTATGATTTTGTGACACACTCAAGTTACCAAGAGACCAAGGATGTCTACCC 486
Oy      486  TCCGACAGCTGCTTCTGTTTGAAGGCACTTGTGTTCTTACACGCCAGGAGATCCGGACAT 545
Db      487  TCCGACAGCTGCTTCTGTTTGAAGGCACTTGTGTTCTTACACGCCAGGAGATCCGGACAT 546
Oy      546  GTTCCAGCTGCGCCCTCTCTGTTGAGACACCGACCTCCGACGTCAGGCTGTCTCCAGAGTCT 605
Db      547  GTTCCAGCTGCGCCCTCTCTGTTGAGACACCGACCTCCGACGTCAGGCTGTCTCCAGAGTCT 606
Oy      606  CCGGACAGCTGCGCCCGAGGAGGAGACCTTGAGAGATTTCTGACGAGTACACACCTTCTGT 665
Db      607  CCGGACAGCTGCGCCCGAGGAGGAGACCTTGAGAGATTTCTGACGAGTACACACCTTCTGT 666
Oy      666  GAAGCCGCGCTTCCAGAGAGTCTTGCC---TGCCGACAAAGAGATATGCCGATGTGATCAT 722
Db      667  GAAGCCGCGCTTCCAGAGAGTCTTGCC---TGCCGACAAAGAGATATGCCGATGTGATCAT 726
Oy      723  CCCACGAGAGTGCACATATGTT 747
Db      727  TCCACGAGAGTGCACATATGANT 751

```

```

RESULT 12
LOCUS   BC025146
DEFINITION Mus musculus uridine monophosphate kinase, mRNA (CDNA clone
ACCESSION BC025146
VERSION   BC025146.1
KEYWORDS   GI:19263563
SOURCE     MGC.
ORGANISM   Mus musculus (house mouse)

```

```

REFERENCE
AUTHORS   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus;
          1 (bases 1 to 1959)
          Strausberg, R.D., Collins, F.S., Wagner, L.H., Derge, J.G.,
          Klausner, R.D., Jordan, H., Moore, T., Max, S.I., Wang, J., Heien, F.,
          Hopkins, R.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
          Datschenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
          Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
          Scheetz, T.E., Brownstein, M.J., Ubedin, T.B., Toshitsuki, S.,
          Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
          Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J.,
          McEwen, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
          Wotley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
          Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
          Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,
          Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
          Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
          Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
          Butlerfield, Y.S., Krzywinski, M.I., Skalska, U., Smalish, D.E.,
          Scherch, A., Schein, J.E., Jones, S.J., and Marra, M.A.
          Generation and initial analysis of more than 15,000 full-length
          human and mouse cDNA sequences
          Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
JOURNAL 22388257
MEDLINE 12477932
PUBMED 2 (bases 1 to 1959)
REFERENCE Strausberg, R.
AUTHORS   Direct Submission
TITLE     Submitted (05-MAR-2002) National Institutes of Health, Mammalian
JOURNAL  Gene Collection (MGC), Cancer Genomics Office, National Cancer
          Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
          USA
REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT   Contact: MGC help desk
          Email: cgaabs-remail.nih.gov
          Tissue Procurement: Jeffrey E. Green, M.D.
          cDNA Library Preparation: Life Technologies, Inc.
          cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

```

DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center
 Center code: BCM-HGSC
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
 Contact: amg@bcm.tmc.edu
 Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulsged, H., Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati, A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAC Plate: 61 Row: p Column: 19
 This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, similarity but not identity to protein.

FEATURES

source

```

1..1959
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="MGC:36231 IMAGE:4913412"
/tissue_type="salivary gland, 10 week old female mouse"
/clone_lib="NCI CGAP_Sc2"
/lab_host="DH10B"
/notes="Vector: pCMV-Sport6"
1..1959
/gene="Umpk"
/db_xref="LocusID:22245"
/db_xref="MGI:98904"
57..890
/codon_start=1
/product="Umpk protein"
/protein_id="AAH25146.1"
/db_xref="GI:19263564"
/db_xref="LocusID:22245"
/translat="MASAGGSGSAAPEADRPORPPFLIGVSGTASGKSTVCEKIM
ELIGQNEVPRORKLVILSDCPYKLTAEQAKLKGQVNPDPAPNDLHMKTLK
NIEGCTVPRPTDPTVTRSRLEPTVYADVLPEGLVFTVQETRMDFHRLPDT
DSVIRSRVLRVQGRDEQLITLYTAFVPAFEFCLPTKADVLIPKVDNMV
AINLIYQHIQDILNGDLCKRHRGPNRHKTFPEPGHPGLATGKSHLESSRP
H"

```

CDS

```

gene
BASE COUNT 466 a 511 c 561 g 421 t
ORIGIN
Query Match 41.0%; Score 666.6; DB 10; Length 1959;
Best Local Similarity 84.4%; Pred. No. 3e-130;
Matches 750; Conservative 0; Mismatches 139; Indels 0; Gaps 0;

```

```

Oy      47  GGGAGGGGGGGGGGGGGGAGCCGATGCGCGGAGACGGCCGAGATGGCTTGGCGG 106
Db      10  GGGCGGAGCGGGCAAGATGCGTAATGATGTCCTGCTCGGATATGAGCTTGGCGG 69
Oy      107 GAGGCGAAGACTGCGAGAGCCCGCCGAGGCGCGACGTCGCGACAGCGGCCCTTCC 166
Db      70  GAGGCGGCGGCTCGAGAGAGCGCGCCGAGGCGCATGTCTCCAGCCGCGCGCTTCC 129
Oy      167 TGATTAGGGGTGAGCGCGCGGACCTGCCAGCGGGAATGCAACCGTGTGTGAAGATCATGG 226
Db      130 TCATCGCGGTGAGACCGCGCGGACCGCTAGTGGCAAGTCAAGTGTGTGAAGATCATGG 189
Oy      227 AGTTGCTGGGAGCAAGACGAGTGGAGACAGCGGAGGGAAGTGTCTGAGGCAAG 286
Db      190 AGCTGCTGGGAGCAAGACGAGTGGAGACAGCGGAGGCAAGTGTGTCTGAGGCAAG 249
Oy      287 ACAGTTTCTTCAAGAGTCTGACGCGAGACAGAGGCGCAAGGCTTGAAGAGACATAGA 346
Db      250 ACAGCTTCTTCAAGAGTCTGACGCGCGAGACAGAGGCGCAAGGCTTGAAGAGACATAGA 309
Oy      347 ATTTTACCATCCAGATGCGCTTTGATTAATGATTGTGACACAGACTCTGAGACATCG 406
Db      310 ATTTTACCATCCAGATGCGCTTTGATTAATGATTGTGACACAGACTCTGAGACATTTG 369

```

QY 407 TGGAGGCAAAACGCTGAGAGTGGCCGACCTATGATTTTGTGACACACTCAAGGTTACAG 466
Db 370 TTGAAGGCAAAACCTGCTGAGAGTCCCTACCTATGATTTTGTGACCACTCAAGGTTACAG 429
QY 467 AGACCAAGGAGTCTACCTGCGGACGCTGCTTCTGTTTGAAGGCACTTGTGTTTACA 526
Db 430 AGACCAAGTGTGCTACCTGCGGACGCTGCTTCTGTTTGAAGGCACTTGTGTTTACA 489
QY 527 GCCAGAGATCCGGGACATGTTCCACTGCGCTCTTCTGTTGACACGACCTCCGAGCTCA 586
Db 490 CCCAGAGATCCGGGACATGTTCCACTGCGCTCTTCTGTTGACACAGACTCTGATGTTA 549
QY 587 GGCCTGTCTGAAAGATCTCCGGGACGCTGCGGACGAGGAGGACCTGAGACATGTTGA 646
Db 550 GGCCTGTCTGAAAGATCTCCGGGACGCTGCGGACGAGGAGGACCTGAGACATGTTGA 609
QY 647 CGGAGTACACACCTTCTGTTGAGGCGGCTTCTGAGGATCTGCTTCCGACAAAGAGT 706
Db 610 CTCAGTACACGCTTCTGTTGAGGCGGCTTCTGAGGATCTGCTTCCGACAAAGAGT 669
QY 707 ATGCCCATGATGATCTCCGACGAGAGTGAACAATATGTTGCTATCACTGATCTGTC 766
Db 670 ACCTGACGATGATCTCCGACGAGAGTGAATAATATGTTGCTATCACTGATCTGTC 729
QY 767 AGCAGATCCAGACATCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 826
Db 730 AACACATCCAGACATCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 789
QY 827 GGGGAGCTTCAAGCGGACCTTTTCTGAGCCAGGAGGACCACTTGGGATCTGATCTG 886
Db 790 GGGGAGCTTCAAGCGGACCTTTTCTGAGCCAGGAGGATCACTTGGGATCTGATCTG 849
QY 887 GCAAGGCTTCAATTTGAGTCCAGCGACGACCCCACTGAGGGGCTGC 935
Db 850 GCAAGGCTTCAATTTGAGTCCAGCGACGACCCCACTGAGGGGCTGC 898

RESULT 13
MUSURKI 1810 bp mRNA linear ROD 27-MAR-1997
LOCUS MUSURKI 1810 bp mRNA linear ROD 27-MAR-1997
DEFINITION Mus musculus uridine kinase mRNA, partial cds.
ACCESSION U11783
VERSION U11783.1 GI:471980
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE 1 (bases 1 to 1810)
AUTHORS Traut, T.W.
TITLE The functions and consensus motifs of nine types of peptide
segments that form different types of nucleotide-binding sites
JOURNAL Eur. J. Biochem. 222 (1), 9-19 (1994)
MEDLINE 94259063
PUBMED 8200357
REFERENCE 2 (bases 1 to 1810)
AUTHORS Kopp, P.A. and Traut, T.W.
TITLE Cloning and expression of a cDNA encoding uridine kinase from mouse
JOURNAL Arch. Biochem. Biophys. 336 (1), 105-112 (1996)
MEDLINE 97108719
PUBMED 8951040
FEATURES
source location/Qualifiers
1..1810
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/tissue_type="brain"
CDS
1..783
/EC_number="2.7.1.48"
/standard_name="uridine-cytidine kinase"
/codon_start=1
/product="uridine kinase"

/protein_id="AAB50568.1"
/db_xref="GI:471981"
/translation="RQPRPPLIGSGTASGKSYCEKIMELLEGONEVDRORKLVI
LSQDCFYKVLAEQAKALKQYINPDPAPFNDDLMHKLKNIIVSKTVEPVYDVT
HSRLPETVTVYPAADVLFEGILVYFQELRDMFHLRLFDLSDVRLSRVLDVORG
RDLEQILTYTAFAVFEFCLPTKQVADVIIPRVDMVVAIINLVQHIQDILNDL
CKRHGGPBNHKKRTPFERGDHPGLATGKSHLESSSRPH"
misc_binding
/bound_molecy="ATP at catalytic site"
polya_signal
BASE COUNT 411 a 482 c 512 g 405 t
ORIGIN
Query Match 38.4%; Score 623; DB 10; Length 1810;
Best Local Similarity 86.7%; Pred. No. 4,8e-121;
Matches 686; Conservative 0; Mismatches 105; Indels 0; Gaps 0;
QY 145 CGTCCGACACGCGGCTTCTGATAGAGGAGCGGCGGCACTCCAGCGGAAAGTCG 204
Db 1 CGTCCGACGCGGCGGCTTCTGATAGAGGAGCGGCGGCACTCCAGCGGAAAGTCG 60
QY 205 ACCGTGTGAGAAATCATGAGTTCTGGGACAGAACGAGGTGGAACGCGGACGG 264
Db 61 ACAGTGTGAGAAATCATGAGTTCTGGGACAGAACGAGGTGGAACGCGGACGGC 120
QY 265 AAGGTGTGATCTGAGCGGACGAGTCTCAAGTCTTCAACGCGGACAGGAGGCG 324
Db 121 AAGGTGTGATCTGAGCGGACGAGTCTTCAAGTCTTCAACGCGGACAGGAGGCG 180
QY 325 AAGGCTTGAAGAGACATATTTTGAACATCCAGATGCTTGTATATGATTTGATG 384
Db 181 AAGGCTTGAAGAGACATATTTTGAACATCCAGATGCTTGTATATGATTTGATG 240
QY 385 CACAGGACTCTGAAGAACTGTGAGGAGCAAAACGTTGAGGTGCGACTATGATTT 444
Db 241 CACAGGACTCTGAAGAACTGTGAGGAGCAAAACGTTGAGGTGCGACTATGATTT 300
QY 445 GTGACACTCAAGTTTACAGAGACCAAGTGTGATACCTTCCGAGCGTGTCTGTT 504
Db 301 GTGACACTCAAGTTTACAGAGACCAAGTGTGATACCTTCCGAGCGTGTCTGTT 360
QY 505 GAGGAGATCTGAGTGTCTTACAGCGGAGATCCGAGATGTTCCAGCTGCGCTCTTC 564
Db 361 GAGGAGATCTGAGTGTCTTACAGCGGAGATCCGAGATGTTCCAGCTGCGCTCTTC 420
QY 565 GTGACACCGACTCCGACGTCAGGCTGTCTGAAAGTCTCCGAGCGTGCAGAGG 624
Db 421 GTGACACCGACTCCGACGTCAGGCTGTCTGAAAGTCTCCGAGCGTGCAGAGG 480
QY 625 AGGAGCTTGAGAGCATTTCTGACGCACTACCACTTGTGAAAGCGGCTTCCAGAG 684
Db 481 AGGAGCTTGAGAGCATTTCTGACGCACTACCACTTGTGAAAGCGGCTTCCAGAG 540
QY 685 TTCTGCTCGGCAAAAGATATGCGGATGATCATCCACGAGAGTGACATATG 744
Db 541 TTCTGCTCGGCAAAAGATATGCGGATGATCATCCACGAGAGTGACATATG 600
QY 745 GTTGCCATCAACCTGATCTGAGACATCCAGACATTTCTGAATGTGACATCTGCAAA 804
Db 601 GTTGCCATCAACCTGATCTGAGACATCCAGACATTTCTGAATGTGACATCTGCAAA 660
QY 805 TGGACCCGAGAGAGGTTCATATGGGCGGAGCTTCAAGCGGACCTTTTCTAGCCAGAGGAC 864
Db 661 TGGACCCGAGAGAGGTTCATATGGGCGGAGCTTCAAGCGGACCTTTTCTAGCCAGAGGAT 720
QY 865 CACCTTGAGATGCTGCTGTGCAAAACGATTCATTTGAGTCCAGAGACGACCCAC 924
Db 721 CACCTTGAGATGCTGCTGTGCAAAACGATTCATTTGAGTCCAGAGACGACCCAC 780
QY 925 TGAAGGCGTGC 935
Db 781 TGAAGGCGTGC 791

```

RESULT 14
MUSURKIA          952 bp   DNA      linear      ROD 05-MAR-2001
LOCUS             Mus musculus uridine kinase pseudogene.
ACCESSION         L31784
VERSION           L31784.1 GI:471982
KEYWORDS          pseudogene; uridine kinase.
SOURCE            Mus musculus (house mouse)
ORGANISM          Mus musculus
                  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE         1 (bases 1 to 952)
AUTHORS           Tsuru, T.W.
TITLE             The functions and consensus motifs of nine types of peptide
                  segments that form different types of nucleotide-binding sites
JOURNAL           Eur. J. Biochem. 222 (1), 9-19 (1994)
MEDLINE           94259063
PUBMED            8200357
REFERENCE         2 (bases 1 to 952)
AUTHORS           Ropp, P.A. and Traut, T.W.
TITLE             Cloning and expression of a cDNA encoding uridine kinase from mouse
                  brain
JOURNAL           Arch. Biochem. Biophys. 336 (1), 105-112 (1996)
MEDLINE           97108719
PUBMED            8951040
FEATURES          source
                  1..952
                  /location="Qualifiers"
                  /organism="Mus musculus"
                  /mol_type="genomic DNA"
                  /db_xref="taxon:10090"
                  /issue_type="kidney"
                  76..913
                  /standard_name="uridine-cytidine kinase"
                  /pseudo
                  /codon_start=1
                  /product="uridine kinase"
                  misc_difference 130
                  /note="insertion not found in L31783"
BASE COUNT       230 a      231 c      277 g      214 t
ORIGIN
Query Match      36.8%; Score 597.4; DB 10; Length 952;
Best Local Similarity 82.3%; Pred. No. 1.2e-115;
Matches 699; Conservative 0; Mismatches 146; Indels 4; Gaps 1;

QY 91 GAGATGGCTTGGGGGGGAGCGAGAGCTGCGAGCCCGC-----GCCGAGGCCGACCG 146
Db 73 GAGATGGCTTGGGGGGGAGCGGAGCTGCTCAGAGAGTACTGCGCTGCCCGAGCGCATCT 132
QY 147 TCCGACACGCGGCGCTTCTGATAGAGGTGAGCGGCGCACTCCAGCGGAGATCGAC 206
Db 133 TCCGACGCGTGGCGGCTTCTCATCGGGGTAGAGGCGGCGACGCGTGTGCGCAAGTCAAC 192
QY 207 CGTGTGAGAAATCATGAGTTGCTGGGACAGAGAGGTGAGCAAGCGGAGCGGAA 266
Db 193 AGTGTGAGAAAGTCAATGAGCTGCTGGGACAGAAAGTGAAGTCACTCCGGTAGCCCAA 252
QY 267 GGTGTATCTCTGAGCGAGGACAGAGTCTCAAGGTCTCTGCGGACAGAGAGGCGCAA 326
Db 253 GTTGTATCTCTGAGCGAGGACGACTCTTCTCAAGGTCTGAGCGGCTAGCGAGAAAGCGAA 312
QY 327 GGCCTTGAAGAGACATATTTTGAACATCCAGATGCCCTTTGATATGATTTGATGCA 386
Db 313 GGCCTTGAAGAGACATATTTTGAACATCCAGATGCCCTTTGATATGATTTGATGCA 372
QY 387 CAGAGCTCTGAAGAAATCGTGAAGGCGAAAACGGTGAAGGTGCGGACCTATGATTTTGT 446
Db 373 CAGAGCTCTGAAGAAACATTTATGAAGGCGAAAGTTGTAAGGTCTCTACTATGATTTTGT 432
QY 447 GACAGACTCAAGGTTTACAGAGACAGGAGGCTTACCTGCGGACGAGTGTCTGTTTGA 506
Db 433 GACGAGCTCAAGGTTTACAGAGACAGGAGGCTTCTTACCAATGATGCGGTTCTGTTTGA 492

```

```

QY 507 GGCATCTTGATGTTTCTACAGAGATCCGGACATGTCACCTGCGCCTTCTGT 566
Db 493 GGCATCTTGATGTTTCTACAGAGATCCGGACATGTTTCTACCTGCGCCTTCTGT 552
QY 567 GAGACCAAGTCTCCAGCTGAGGCTGTCTCGAAGAGTTCTCCGGAGCTGCGGAGGAG 626
Db 553 GAGACCAAGTCTCGATGATTTAGGCTGTCTTGAAGGTTCTCCGGAGTGTCAACGAGGAG 612
QY 627 GAGCTTGAGACATTTCTGACGATACACACCTTCTGTAAACCGGCTTCCAGAGATT 686
Db 613 GAGCTTGAGACATTTCTGACGATACACACCTTCTGTAAACCGGCTTCCAGAGATT 672
QY 687 CTGCTCCGACAAAGAGATGATCCGATGATCATCCAGAGATGAGCAATATGTT 746
Db 673 CTGCTCCGACAAAGAGATGATCCGATGATCATCCAGAGATGAGCAATATGTT 732
QY 747 TGCATCACTGATGCTGAGACATCCAGACATTTGATGATGATGATGATGATG 806
Db 733 GGCATCACTGATGCTGATGATGATGATGATGATGATGATGATGATGATGATG 792
QY 807 GACCCGAGGAGGATCCAAATGGGAGGCTCAAGCGGACCTTTTGTAGCGAGGAGCA 866
Db 793 GACCCGAGGAGGATCCAAATGGGAGGCTCAAGCGGACCTTTTGTAGCGAGGAGCA 852
QY 867 CCTGGAGTCTGACCTTGGCAAGCGGTGATGATGATGATGATGATGATGATGATG 926
Db 853 CCTGGAGTCTTGGCAAGCGGTGATGATGATGATGATGATGATGATGATGATGATG 912
QY 927 AGGAGCTGC 935
Db 913 AGGAGCTGC 921

```

```

RESULT 15
AC078885/c      254993 bp   DNA      linear      HTG 15-MAY-2002
LOCUS           Mus musculus chromosome 6 clone RP23-117123 strain C57BL/6J,
DEFINITION      WORKING DRAFT SEQUENCE, 44 unordered pieces.
ACCESSION       AC078885
VERSION         AC078885.8 GI:18376842
KEYWORDS        HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE          Mus musculus (house mouse)
ORGANISM        Mus musculus
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE       1 (bases 1 to 254993)
AUTHORS         Montgomery, K.T., Grills, G., Han, J., Lee, E., Long, J., Pomerantz, R.,
                Ioshikhes, I.P., Shim, C., Decker, J., Thomas, E., Pereira, A.,
                Gordon, M., Goltz, J.S. and Kucherlapati, R.
TITLE           High Throughput Mouse Sequencing
JOURNAL         Unpublished
AUTHORS         2 (bases 1 to 254993)
                Montgomery, K.T., Grills, G., Han, J., Lee, E., Long, J., Pomerantz, R.,
                Ioshikhes, I.P., Shim, C., Decker, J., Thomas, E., Pereira, A.,
                Gordon, M., Goltz, J.S. and Kucherlapati, R.
TITLE           Direct Submission
JOURNAL         Submitted (08-NOV-2000) Department of Molecular Genetics, Albert
                Einstein College of Medicine Genome Center, 1300 Morris Park Ave.,
                Bronx, NY 10461, USA
                On Jan 26, 2002 this sequence version replaced gi:14488282.
COMMENT         -----Genome Center
                Center: Harvard Partners Genome Center
                Center Code: HPGC
                Web site: http://www.hpcgg.org/Sequence/mouse.html
                Contact: hpcgcmendel.mgh.harvard.edu
                -----Summary Statistics
                Center Project name: AAF
                Sequencing vector: pUC18; L08752
                Chemistry: Dye-terminator Big Dye 1.00%
                *Consensus quality: 247096 at least 020
                *Consensus quality: 242518 at least 040
                *Consensus quality: 235589 at least Q30

```

```
Estimated insert size: agarose-FP - N/A
**Estimated insert size: 254133 - sum-of-contigs
Quality coverage: agarose-FP - N/A
Quality coverage: 6 x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 44 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
27415: contig of 27415 bp in length
27416 27435: gap of unknown length
27436 57817: contig of 30382 bp in length
57818 57837: gap of unknown length
57838 81637: contig of 23800 bp in length
81638 81657: gap of unknown length
81658 102155: contig of 20498 bp in length
102156 102175: gap of unknown length
102176 121200: contig of 19025 bp in length
121201 121220: gap of unknown length
121221 135073: contig of 13853 bp in length
135074 135093: gap of unknown length
135094 156688: contig of 21596 bp in length
156689 156703: gap of unknown length
156704 166185: contig of 9476 bp in length
166186 166205: gap of unknown length
166206 178798: contig of 12593 bp in length
178799 178818: gap of unknown length
178819 187247: contig of 8429 bp in length
187248 187267: gap of unknown length
187268 192492: contig of 5225 bp in length
192493 192512: gap of unknown length
192513 200084: contig of 7572 bp in length
200085 200104: gap of unknown length
200105 207054: contig of 6950 bp in length
207055 207074: gap of unknown length
207075 212944: contig of 5870 bp in length
212945 212964: gap of unknown length
212965 219468: contig of 6504 bp in length
219469 219488: gap of unknown length
219489 220369: contig of 881 bp in length
220370 220389: gap of unknown length
220390 221164: contig of 775 bp in length
221165 221184: gap of unknown length
221185 222796: contig of 1612 bp in length
222797 222816: gap of unknown length
222817 224762: contig of 1946 bp in length
224763 224782: gap of unknown length
224783 225999: contig of 1217 bp in length
225999 226019: gap of unknown length
226020 227409: contig of 1390 bp in length
227410 227429: gap of unknown length
227430 230145: contig of 2716 bp in length
230146 230165: gap of unknown length
230166 231842: contig of 1677 bp in length
231843 231862: gap of unknown length
231863 232786: contig of 924 bp in length
232787 232806: gap of unknown length
232807 233137: contig of 331 bp in length
233138 233157: gap of unknown length
233158 234519: contig of 1362 bp in length
234519 234539: gap of unknown length
234540 236416: contig of 1877 bp in length
236417 236436: gap of unknown length
236437 237967: contig of 1531 bp in length
237968 237987: gap of unknown length
237988 239220: contig of 1233 bp in length
239221 239240: gap of unknown length
239241 239367: contig of 127 bp in length
239368 239387: gap of unknown length

FEATURES
source
239388 239959: contig of 572 bp in length
239960 239979: gap of unknown length
239980 241016: contig of 1037 bp in length
241017 241036: gap of unknown length
241037 242391: contig of 1355 bp in length
242392 242391: gap of unknown length
242392 242411: gap of unknown length
242412 243823: contig of 1412 bp in length
243824 243843: gap of unknown length
243844 245227: contig of 1384 bp in length
245228 245247: gap of unknown length
245248 246334: contig of 1087 bp in length
246335 246355: gap of unknown length
246355 247451: contig of 1097 bp in length
247452 247471: gap of unknown length
247472 248780: contig of 1309 bp in length
248781 248800: gap of unknown length
248801 250041: contig of 1241 bp in length
250042 250061: gap of unknown length
250062 251173: contig of 1112 bp in length
251174 251193: gap of unknown length
251194 252713: contig of 1520 bp in length
252714 252733: gap of unknown length
252734 254015: contig of 1281 bp in length
254015 254034: gap of unknown length
254035 254176: contig of 142 bp in length
254177 254196: gap of unknown length
254197 254993: contig of 797 bp in length.
Location/Qualifiers
1. 254993
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL6/J"
/db_xref="taxon:10090"
/chromosome="6"
/clone="RP23-117123"
/sex="male"
1. 27415
/feature="assembly_name:Contig138"
27436. 57817
/feature="assembly_name:Contig137
clone_end:Sp6
vector_side:left"
57818. 81637
/feature="assembly_name:Contig136"
81658. 102155
/feature="assembly_name:Contig135"
102176. 121200
/feature="assembly_name:Contig134"
121221. 135073
/feature="assembly_name:Contig133"
135094. 156689
/feature="assembly_name:Contig132"
156710. 166185
/feature="assembly_name:Contig131
clone_end:T7
vector_side:left"
166206. 178798
/feature="assembly_name:Contig130"
178819. 187247
/feature="assembly_name:Contig129"
187268. 192492
/feature="assembly_name:Contig128"
192513. 200084
/feature="assembly_name:Contig127"
200105. 207054
/feature="assembly_name:Contig126"
207075. 212944
/feature="assembly_name:Contig125"
212965. 219468
/feature="assembly_name:Contig124"
219489. 220369
/feature="assembly_name:Contig123"
220390. 221164
```

```
misc_feature 221185..222796 /note="assembly_name:Contig122"  
misc_feature 222817..224762 /note="assembly_name:Contig121"  
misc_feature 224783..225399 /note="assembly_name:Contig120"  
misc_feature 226020..227409 /note="assembly_name:Contig119"  
misc_feature 227430..230145 /note="assembly_name:Contig118"  
misc_feature 230166..231842 /note="assembly_name:Contig117"  
misc_feature 231863..232786 /note="assembly_name:Contig116"  
misc_feature 232807..233137 /note="assembly_name:Contig115"  
misc_feature 233158..233487 /note="assembly_name:Contig114"
```

Query Match	35.6%;	Score 594.6;	DB 2;	Length 254933;
Best Local Similarity	79.8%;	Pred. No. 4.6e-115;		
Matches 700;	Conservative	0;	Mismatches 174;	Indels 3; Gaps 1;

[illegible]

QY	928	G G G G T G C C A G A G C C T C A G S G C A G H C T C C C G C C G A T G T G N T T C A G S A C T G A G C C T	987
Db	155935	G G A C C A G A T A --- T G T A G T T C C C C A C A G A C C A G A T T A A G G G C C T G G G A T G A C A C	155879
QY	968	G G G A C G C C C A C C C A C A C C C A C T G C T T C T C T C G S G	1024
Db	155878	T G G C T C T A A A A G C A C A G T C A G G A T C C C T C T C A G G	155842

Search completed: November 25, 2003, 00:21:58
Job time : 6324.34 secs

QY 181 GCGGCGACTGCGCAGCGGGAAGTCGACCGCTGTGTGAGAGATCATGAGTTGCTGGGACAG 240
DB 181 GCGGCGACTGCGCAGCGGGAAGTCGACCGCTGTGTGAGAGATCATGAGTTGCTGGGACAG 240
QY 241 AACGAGGTGGAACAGCGGCGAGCGGAAGTGGTCATCTTGAGCCGAGACAGATTCTACAG 300
DB 241 AACGAGGTGGAACAGCGGCGAGCGGAAGTGGTCATCTTGAGCCGAGACAGATTCTACAG 300
QY 301 GTCCTGACGCGAGAGAGAGAGCGCAAGGCTTTGAAGAGACATGAAATTTTACCATCA 360
DB 301 GTCCTGACGCGAGAGAGAGAGCGCAAGGCTTTGAAGAGACATGAAATTTTACCATCA 360
QY 361 GATGCGCTTTGATATGATTTGATGACAGAGCTGTGAAGAACATCGTGAGGGGCAAAACG 420
DB 361 GATGCGCTTTGATATGATTTGATGACAGAGCTGTGAAGAACATCGTGAGGGGCAAAACG 420
QY 421 GTGAGAGTGCAGACTATGATTTTGTGACACACTCAAGTTTACGAGACCAAGGTTGTC 480
DB 421 GTGAGAGTGCAGACTATGATTTTGTGACACACTCAAGTTTACGAGACCAAGGTTGTC 480
QY 481 TACCTGCGGAGCGTGTCTGTGTGAGAGGACATCTTGATTTTACAGCCAGAGATCCGG 540
DB 481 TACCTGCGGAGCGTGTCTGTGTGAGAGGACATCTTGATTTTACAGCCAGAGATCCGG 540
QY 541 GACATGTTTCAACCTGCGGCTTCTGTGAGACACGAGCTCCGACGCTGCTCCGAGA 600
DB 541 GACATGTTTCAACCTGCGGCTTCTGTGAGACACGAGCTCCGACGCTGCTCCGAGA 600
QY 601 GTTCTCCGAGACGTCGCCCGAGAGAGGACCTGTGAGAGATTTGAGCGAGTACACCA 660
DB 601 GTTCTCCGAGACGTCGCCCGAGAGAGGACCTGTGAGAGATTTGAGCGAGTACACCA 660
QY 661 TTGCTGAAGCGGCTTCTGAGAGATTTGCTGCTGCGACAAAGATATGCCGATGTATC 720
DB 661 TTGCTGAAGCGGCTTCTGAGAGATTTGCTGCTGCGACAAAGATATGCCGATGTATC 720
QY 721 ATCCGAGAGAGAGGACAAATATGTTGTCATGAACTGATCGTGAGACATCCAGAGAC 780
DB 721 ATCCGAGAGAGAGGACAAATATGTTGTCATGAACTGATCGTGAGACATCCAGAGAC 780
QY 781 ATTCTGATGATGATCATCTGCAATATGAGACCGAGAGAGGTTCAATGGGCGAGCTACAG 840
DB 781 ATTCTGATGATGATCATCTGCAATATGAGACCGAGAGAGGTTCAATGGGCGAGCTACAG 840
QY 841 CGGACCTTTTCTGAGCCAGGAGACCACTGAGGATGCTGACCTTGCCAAACGGTCAAT 900
DB 841 CGGACCTTTTCTGAGCCAGGAGACCACTGAGGATGCTGACCTTGCCAAACGGTCAAT 900
QY 901 TTGAGATCCAGACAGACCCCACTGAGGAGGCTGCGAGGCTGAGGCGAGGTCCTCCGCC 960
DB 901 TTGAGATCCAGACAGACCCCACTGAGGAGGCTGCGAGGCTGAGGCGAGGTCCTCCGCC 960
QY 961 CGGAGATGTTGTTGAGGAGCTGAGCTGAGGAGCGCCACCACTGAGCTGCTCTC 1020
DB 961 CGGAGATGTTGTTGAGGAGCTGAGCTGAGGAGCGCCACCACTGAGCTGCTCTC 1020
QY 1021 GGGGCACTCCAGAGGAGTGTAGAGAGGCTTTCTCACTCAGAGTGGAAAATCTAGA 1080
DB 1021 GGGGCACTCCAGAGGAGTGTAGAGAGGCTTTCTCACTCAGAGTGGAAAATCTAGA 1080
QY 1081 TGTGTCACTCAGACTCAACTGTGAGGACCTGACAGCGCTTCTGAGTTTTCAGCCAC 1140
DB 1081 TGTGTCACTCAGACTCAACTGTGAGGACCTGACAGCGCTTCTGAGTTTTCAGCCAC 1140
QY 1141 TTAGGCTGTTGCGGTTTAAAGATCCCTTATGATCTGAGAAATGCAAGATGTGCA 1200
DB 1141 TTAGGCTGTTGCGGTTTAAAGATCCCTTATGATCTGAGAAATGCAAGATGTGCA 1200
QY 1201 GGAAGCTTGGAGGCTTCTGTGAGGAAATGTAGGCACTTATTTGGGAAAATGAGAGAC 1260
DB 1201 GGAAGCTTGGAGGCTTCTGTGAGGAAATGTAGGCACTTATTTGGGAAAATGAGAGAC 1260

QY 1261 AGCTTAGACACTGCGGCTGATGTTGTTGATGAGTGAACCCAGTGGGAGAGATT 1320
DB 1261 AGCTTAGACACTGCGGCTGATGTTGTTGATGAGTGAACCCAGTGGGAGAGATT 1320
QY 1321 TTTTCCAGTCTGATCTGTTCTTTACACTCACTCAACATTAATCAAAAGTTTGTGACA 1380
DB 1321 TTTTCCAGTCTGATCTGTTCTTTACACTCACTCAACATTAATCAAAAGTTTGTGACA 1380
QY 1381 AGTACTTCTCTTTTACATGTTACATGTCCTCAGTTTCTGTTTCTGTTTCAATAA 1440
DB 1381 AGTACTTCTCTTTTACATGTTACATGTCCTCAGTTTCTGTTTCTGTTTCAATAA 1440
QY 1441 CAAGGCTGTTGAGGCTTCAAACTTAATTTATGACCCAGTGTGAGTCCAGCGTG 1500
DB 1441 CAAGGCTGTTGAGGCTTCAAACTTAATTTATGACCCAGTGTGAGTCCAGCGTG 1500
QY 1501 GCCTTACCGGATATGAGGAGCCACTGAGGAGATGTTTCCCTTGTGCTTTAAG 1560
DB 1501 GCCTTACCGGATATGAGGAGCCACTGAGGAGATGTTTCCCTTGTGCTTTAAG 1560
QY 1561 GCAGAGAACGAGGCGGATGCCCTGAGACACCCAGATCAACCAAGCTTGTGGGG 1620
DB 1561 GCAGAGAACGAGGCGGATGCCCTGAGACACCCAGATCAACCAAGCTTGTGGGG 1620
QY 1621 CCAG 1624
DB 1621 CCAG 1624

RESULT 2
US-09-833-381-2048
; Sequence 2048 Application US/09833381
; Patent No. US20020132090A1
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. US20020132090A1 Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; PRIORITY FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2048
; LENGTH: 1648
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(1648)
; OTHER INFORMATION: n = A,T,C or G
US-09-833-381-2048

Query Match 86.9%; Score 1411; DB 10; Length 1648;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 46 GGGGAAAGGCGGCGCGGAGACCCGATGCGGAGCGGAGCGGAGCGGAGTGGCTTGGCGG 105
DB 36 GGGGAAAGGCGGCGCGGAGACCCGATGCGGAGCGGAGCGGAGCGGAGTGGCTTGGCGG 95
QY 106 GGAAGGGAAGACTGCGAGAGCCCGCGGAGGCGGAGCGGAGCGGAGCGGAGCGGAGCGGCTTC 165
DB 96 GGAAGGGAAGACTGCGAGAGCCCGCGGAGGCGGAGCGGAGCGGAGCGGAGCGGAGCGGCTTC 155
QY 166 CTGATAGGGGTGAGCGGCGGCACTGCGACGCGGAAAGTCAACCGTGTGTGAGAAATCATG 225
DB 156 CTGATAGGGGTGAGCGGCGGCACTGCGACGCGGAAAGTCAACCGTGTGTGAGAAATCATG 215
QY 226 GAATTCTGAGACAGAACGAGTGTGAACAGCGGACGGAAGGTGTCTTCTGAGCCAG 285
DB 216 GAATTCTGAGACAGAACGAGTGTGAACAGCGGACGGAAGGTGTCTTCTGAGCCAG 275

QY 286 GACAGGTTCTACAGGTCCTGACGGCAGACAGAAAGCCAAAGCCTTGAAGAAGACAGTAC 345
| | | | |
Db 276 GACAGGTTCTACAGGTCCTGACGGCAGACAGAAAGCCAAAGCCTTGAAGAAGACAGTAC 335
| | | | |
QY 346 AATTTTGAACCATTCAGATGCTTTTGAATATGATTTGATGACAGAGACTTGAAAGACATC 405
| | | | |
Db 336 AATTTTGAACCATTCAGATGCTTTTGAATATGATTTGATGACAGAGACTTGAAAGACATC 395
| | | | |
QY 406 GTGAGAGGGCAAAACGGTGGAGGTGCCGACTATGATTTTGTGACACCTCAAGATTACCA 465
| | | | |
Db 396 GTGAGAGGGCAAAACGGTGGAGGTGCCGACTATGATTTTGTGACACCTCAAGATTACCA 455
| | | | |
QY 466 GAGACCAAGGTGTCTACCTTCGGGAGGTGCTTGTGTTGAGGAGCATCTTGATGTTCTAC 525
| | | | |
Db 456 GAGACCAAGGTGTCTACCTTCGGGAGGTGCTTGTGTTGAGGAGCATCTTGATGTTCTAC 515
| | | | |
QY 526 AGCCAGAGATCCGGGACATGTTTCAACTGCGCTCTTGTGTTGACACCGACTCCGAGCTC 585
| | | | |
Db 516 AGCCAGAGATCCGGGACATGTTTCAACTGCGCTCTTGTGTTGACACCGACTCCGAGCTC 575
| | | | |
QY 586 AGGCTGTCTCGAAGAGTCTCCGGGAGGTGCGCCGAGGAGGAGCCTGAGAGCATTTCTG 645
| | | | |
Db 576 AGGCTGTCTCGAAGAGTCTCCGGGAGGTGCGCCGAGGAGGAGCCTGAGAGCATTTCTG 635
| | | | |
QY 646 AGCAGATACACCACTTCGTGAAGCCGGCTTGAAGAGTTCGTGCTGCCGAGAAAGAG 705
| | | | |
Db 636 AGCAGATACACCACTTCGTGAAGCCGGCTTGAAGAGTTCGTGCTGCCGAGAAAGAG 695
| | | | |
QY 706 TA-TGCCGATGTATCATCCACGAGAGGTGACAAATATGTTGCCATCAACTGATCGT 764
| | | | |
Db 696 TATTCGCGATGTATCATCCACGAGAGGTGACAAATATGTTGCCATCAACTGATCGT 755
| | | | |
QY 765 GCAGCATCTCAGAGCATTTGAAATGTGTGCAATCTGCAAAATGGCAGGAGGAGGTCCAA 824
| | | | |
Db 756 GCAGCATCTCAGAGCATTTGAAATGTGTGCAATCTGCAAAATGGCAGGAGGAGGTCCAA 815
| | | | |
QY 825 TGGGCGAGGCTACAGAGGACCTTTTCTGAGCAGGAGGACCACTCGTGGATGTCGACCTC 884
| | | | |
Db 816 TGGGCGAGGCTACAGAGGACCTTTTCTGAGCAGGAGGACCACTCGTGGATGTCGACCTC 875
| | | | |
QY 885 TGGCAAAACGGTCAATTTGGAGTTCAGCAGACAGACCCCACTGAAGGGCTGCCAGCTCA 944
| | | | |
Db 876 TGGCAAAACGGTCAATTTGGAGTTCAGCAGACAGACCCCACTGAAGGGCTGCCAGCTCA 935
| | | | |
QY 945 GGGCAGAGTCTCCCGCCCGAGATGTGTGTTAGAGGACTGAGCCTGGGAGCCCAACCCACA 1004
| | | | |
Db 936 GGGCAGAGTCTCCCGCCCGAGATGTGTGTTAGAGGACTGAGCCTGGGAGCCCAACCCACA 995
| | | | |
QY 1005 CCCACTGCTTCTCTCGGCGCACCCCAAGGGAGGTGTTAGCAGGAGGCTTCTCACTCA 1064
| | | | |
Db 996 CCCACTGCTTCTCTCGGCGCACCCCAAGGGAGGTGTTAGCAGGAGGCTTCTCACTCA 1055
| | | | |
QY 1065 GGAGTGAAGAACTCAGATGTGTCTCACTCAGACTCACTTGCGGACACTGACAGGCTTCC 1124
| | | | |
Db 1056 GGAGTGAAGAACTCAGATGTGTCTCACTCAGACTCACTTGCGGACACTGACAGGCTTCC 1115
| | | | |
QY 1125 TGAAGTTTTCAGCAGCTTGAAGGCTGTTGCGGTTTAAAGATCCCTCTAAGTCACTGAGAAA 1184
| | | | |
Db 1116 TGAAGTTTTCAGCAGCTTGAAGGCTGTTGCGGTTTAAAGATCCCTCTAAGTCACTGAGAAA 1175
| | | | |
QY 1185 TGGCACAAGATGTCAGAGAAAGCTTGAGAGGCTTCTGAGAGGAATGAGGCAATTATTG 1244
| | | | |
Db 1176 TGGCACAAGATGTCAGAGAAAGCTTGAGAGGCTTCTGAGAGGAATGAGGCAATTATTG 1235
| | | | |
QY 1245 GGGGAAATGAGAGACAGCTTAGACACTGGCTGGCCTGATGTTTGTGTTGACATGAGACCC 1304
| | | | |
Db 1236 GGGGAAATGAGAGACAGCTTAGACACTGGCTGGCCTGATGTTTGTGTTGACATGAGACCC 1295
| | | | |
QY 1305 ACAGTGGAGAGAGTTTTCAGTCTGATCTGTTCTTACACTGACACTGACATTAATCTC 1364
| | | | |
Db 1296 ACAGTGGAGAGAGTTTTCAGTCTGATCTGTTCTTACACTGACACTGACATTAATCTC 1355
| | | | |
QY 1365 AAAAGTTTGTGAACAAGTACTTCTTTTATCATGTTACATGATCTGATGTTTCTGTG 1424
| | | | |

Db 1356 AAAAGTTTGTGAACAAGTACTTCTTTTATCATGTTACATGTCATGTTTCTGT 1415
| | | | |
QY 1425 TTTCTGTTCAATPACCAAGGCTGTTGTGCGCTTACAAACCTTAATTTCAATGACCCAGTGG 1484
| | | | |
Db 1416 TTTCTGTTCAATPACCAAGGCTGTTGTGCGCTTACAAACCTTAATTTCAATGACCCAGTGG 1475
| | | | |
QY 1485 TTTGCACTCAGCGTGGCCCTTACACGGATATGAGGAGCCCACTGAGGGAATGTTTCCCGCTC 1544
| | | | |
Db 1476 TTTGCACTCAGCGTGGCCCTTACACGGATATGAGGAGCCCACTGAGGGAATGTTTCCCGCTC 1535
| | | | |
QY 1545 TGCCTGTGCTTAAAGGACAGAGAGGCGG 1576
| | | | |
Db 1536 TGCCTGTGCTTAAAGGACAGAGAGGCGG 1567
| | | | |
RESULT 3
US-10-037-270-546
; Sequence 546, Application US/10037270
; Publication No. US20030104529A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonphong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yungting
; APPLICANT: Wang, Duntui
; APPLICANT: Wang, Zhiwei
; APPLICANT: Tillinphast, John
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20030104529A1el Nucleic Acids and
; FILE REFERENCE: Polypeptides
; CURRENT APPLICATION NUMBER: US/10/037,270
; PRIOR FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: pt FL_genes Version 1.0
; SEQ ID NO 546
; LENGTH: 1288
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (95) .. (928)
US-10-037-270-546
Query Match 57.4%; Score 932; DB 14; Length 1288;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 992; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 46 GGGGAAAGGGCGGGCGCGGGGACCCGATGCGCGGAGCGGAGCGGAGATGCTTCGGCG 105
| | | | |
Db 47 GGGGAAAGGGCGGGCGCGGGGACCCGATGCGCGGAGCGGAGCGGAGATGCTTCGGCG 106
| | | | |
QY 106 GGAAGCGGAAGCTGGAGAGCCCGCGCGGAGGCGGACCGTCCGACACCAAGGCGGCTTC 165
| | | | |
Db 107 GGAAGCGGAAGCTGGAGAGCCCGCGCGGAGGCGGACCGTCCGACACCAAGGCGGCTTC 166
| | | | |
QY 166 CTGATAGGGGTAGGCGCGGCACTGCAAGCGGAAATGCAACCGTGTGTGAAGAATCATG 225
| | | | |
Db 167 CTGATAGGGGTAGGCGCGGCACTGCAAGCGGAAATGCAACCGTGTGTGAAGAATCATG 226
| | | | |

```
QY 226 GAGTTGCTGGGACGAAACGAGTGGAAACAGCGGACGCGGAAGTGTGATCTTGAGCCAG 285
|||
DB 227 GAGTGTCTGGGACGAAACGAGTGGAAACAGCGGACGCGGAAGTGTGATCTTGAGCCAG 286
286 GACAGGTTCTACAGAGTCTGACGCGACAGACGAAAGGCTTGAAAGGACAGTAC 345
287 GACAGGTTCTACAGAGTCTGACGCGACAGACGAAAGGCTTGAAAGGACAGTAC 346
QY 346 AATTTTACCATCCAGATGCTTTTGAATGATTTGATGACACAGAGCTTGAAAGACATC 405
|||
DB 347 AATTTTACCATCCAGATGCTTTTGAATGATTTGATGACACAGAGCTTGAAAGACATC 406
QY 406 GTGAGGCGCAAAACGGTGGAGGTGCGACCTATATATTTTGTGACACATCAAGTTTACA 465
|||
DB 407 GTGAGGCGCAAAACGGTGGAGGTGCGACCTATATATTTTGTGACACATCAAGTTTACA 466
466 GAGACACAGTGTCTACCTGCGGACGATGCTTGTGTTGAGGACATCTTGTTTCTAC 525
|||
DB 467 GAGACACAGTGTCTACCTGCGGACGATGCTTGTGTTGAGGACATCTTGTTTCTAC 526
QY 526 AGCCAGAGATCCGGGACATGTTCCACTGCGCTCTTCTGTGACACCGACTCCGACGTC 585
|||
DB 527 AGCCAGAGATCCGGGACATGTTCCACTGCGCTCTTCTGTGACACCGACTCCGACGTC 586
586 AGGCTGTCTGGAAGATGTTCCCGGGACGTGCGCGAGGAGGACCTGAGACAGATTTCTG 645
|||
DB 587 AGGCTGTCTGGAAGATGTTCCCGGGACGTGCGCGAGGAGGACCTGAGACAGATTTCTG 646
QY 646 ACAGCAGTACACACCTTCTGTGAAACCGGCTTTCAGAGAGTCTGCTGCGACCAAGAAAG 705
|||
DB 647 ACAGCAGTACACACCTTCTGTGAAACCGGCTTTCAGAGAGTCTGCTGCGACCAAGAAAG 706
QY 706 TATGCCATGTGATATCCACGAGAGTGGACAATATGTTGCTCATCACTGATGTG 765
|||
DB 707 TATGCCATGTGATATCCACGAGAGTGGACAATATGTTGCTCATCACTGATGTG 766
QY 766 CAGCAGATCCAGGACATTTCTGAAATGATGACATCTGCAAAATGGACCGAGAGGCTCCAT 825
|||
DB 767 CAGCAGATCCAGGACATTTCTGAAATGATGACATCTGCAAAATGGACCGAGAGGCTCCAT 826
QY 826 GGGGAGAGCTACAGCGGACCTTTTCTGAGCCAGGGGACCACTTGAGGATGCTGACCTCT 885
|||
DB 827 GGGGAGAGCTACAGCGGACCTTTTCTGAGCCAGGGGACCACTTGAGGATGCTGACCTCT 886
QY 886 GGCAGACGCTACATTTGAGAGTCCAGACGACCCCACTGAGGGGCTGCCGAGCTCAG 945
|||
DB 887 GGCAGACGCTACATTTGAGAGTCCAGACGACCCCACTGAGGGGCTGCCGAGCTCAG 946
QY 946 GGCAGGTCCTCCGCGCGGACATGTTGTTCAAG 977
|||
DB 947 GGCAGGTCCTCCGCGCGGACATGTTGTTCAAG 978

RESULT 4
US-09-764-877-2804
; Sequence 2804, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005
; CURRENT APPLICATION NUMBER: US/09/764,877
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: Patent Ver. 2.0
; SEQ ID NO 2804
; LENGTH: 9732
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-877-2804
```

```
Query Match 54.2%; Score 880; DB 10; Length 9732;
Best Local Similarity 100.0%; Pred. No. 0; Indels 0; Gaps 0;
Matches 880; Conservative 0; Mismatches 0;

QY 745 GTTGCCATCAACCTGATCTGTGACGACATCCAGACATTTCTGAATGTGACATCTGCAAA 804
|||
DB 3901 GTTGCCATCAACCTGATCTGTGACGACATCCAGACATTTCTGAATGTGACATCTGCAAA 3960
QY 805 TGGACCCGAGAGAGGTTCCAAATGGGCGGAGCTACAAAGGACCTTTTCTAGCCAGGGGAC 864
|||
DB 3961 TGGACCCGAGAGAGGTTCCAAATGGGCGGAGCTACAAAGGACCTTTTCTAGCCAGGGGAC 4020
QY 865 CACCTGGATGTGACCTCTGACCAACGATCAATTTGAGTCCAGACAGACCCAC 924
|||
DB 4021 CACCTGGATGTGACCTCTGACCAACGATCAATTTGAGTCCAGACAGACCCAC 4080
QY 925 TGAAGGAGCTGCCGAGCTCAGAGGAGTCTCCCGCGGACATGTGTTCAGGACCTGAG 984
|||
DB 4081 TGAAGGAGCTGCCGAGCTCAGAGGAGTCTCCCGCGGACATGTGTTCAGGACCTGAG 4140
QY 985 CTTGGGAGAGCCCAACCAACCACTGCTTCTTCTGGGACCCCAAGGGAGTGTAC 1044
|||
DB 4141 CTTGGGAGAGCCCAACCAACCACTGCTTCTTCTGGGACCCCAAGGGAGTGTAC 4200
QY 1045 AGCAGAGCTTCTCTCACTCAGAGTGGAACTCAGATGTGCACTCAGACTCAACTTGCT 1104
|||
DB 4201 AGCAGAGCTTCTCTCACTCAGAGTGGAACTCAGATGTGCACTCAGACTCAACTTGCT 4260
QY 1105 GGGACACTGACAGAGCTTCTGAGGTTTCAAGCACTTAAGCTCTGTGCTTAAAGAT 1164
|||
DB 4261 GGGACACTGACAGAGCTTCTGAGGTTTCAAGCACTTAAGCTCTGTGCTTAAAGAT 4320
QY 1165 CCTCTAGTCACTGAGAAATGCCACAGATGTGACAGAAAGCTGGAGGCTTCTGTGAG 1224
|||
DB 4321 CCTCTAGTCACTGAGAAATGCCACAGATGTGACAGAAAGCTGGAGGCTTCTGTGAG 4380
QY 1225 GAATGTGAGGACATTAATGGGAAATGAGGAGACAGCTTACACATGCTGCTGAT 1284
|||
DB 4381 GAATGTGAGGACATTAATGGGAAATGAGGAGACAGCTTACACATGCTGCTGAT 4440
QY 1285 GTTTTGTGACAGTGAACCAACAGTGGAGAGATTTTTCAGTCTGATCTGCTTAA 1344
|||
DB 4441 GTTTTGTGACAGTGAACCAACAGTGGAGAGATTTTTCAGTCTGATCTGCTTAA 4500
QY 1345 CACACTCACACATTAATCTCAAAAGTTTGTGAACAGTACTTCTTTTAAACAGTAA 1404
|||
DB 4501 CACACTCACACATTAATCTCAAAAGTTTGTGAACAGTACTTCTTTTAAACAGTAA 4560
QY 1405 CATGTCCTCATGTTTCTGTTTCTGTTTCTATTAACAAGGCTGTTGTGCTTAAAC 1464
|||
DB 4561 CATGTCCTCATGTTTCTGTTTCTGTTTCTATTAACAAGGCTGTTGTGCTTAAAC 4620
QY 1465 CTAAATTCATGACCCAGTGGTTTTCAGTCCAGGCTGCTTACACGATATGGGAGCCAC 1524
|||
DB 4621 CTAAATTCATGACCCAGTGGTTTTCAGTCCAGGCTGCTTACACGATATGGGAGCCAC 4680
QY 1525 TGAAGGATGTTTCCCGCTGCTTGAAGGAGAGAGGAGGAGGAGGAGGAGGAGGAG 1584
|||
DB 4681 TGAAGGATGTTTCCCGCTGCTTGAAGGAGAGAGGAGGAGGAGGAGGAGGAGGAG 4740
QY 1585 GGAAGACCCAGCATCACACCCAGGCTTGGCGGGCCAG 1624
|||
DB 4741 GGAAGACCCAGCATCACACCCAGGCTTGGCGGGCCAG 4780

RESULT 5
US-09-764-877-2805
; Sequence 2805, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005
```

```
/ CURRENT APPLICATION NUMBER: US/09/764,877
/ CURRENT FILING DATE: 2001-01-17
/ Prior application data removed - refer to PALM or file wrapper
/ NUMBER OF SEQ ID NOS: 4031
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 2805
/ LENGTH: 19125
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-764-877-2805

Query Match      54.2%; Score 880; DB 10; Length 19125;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 880; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 745 GTTCCATCACTGATGTCGACGACATCCAGACATTTGTAATGTCATCTGCAA 804
DB 5997 GTTGCCATCACTGATGTCGACGACATCCAGACATTTGTAATGTCATCTGCAA 6056
QY 805 TGGACCCGAGAGAGGTCCTCAATGGGCGGAGCTACAGCGGACCTTTCTGAGCGAGGAG 864
DB 6057 TGGACCCGAGAGAGGTCCTCAATGGGCGGAGCTACAGCGGACCTTTCTGAGCGAGGAG 6116
QY 865 CACCCTGGATGTCGACCTCTGGCAAACGGTCAATTGGAGTCCAGACAGACCCAC 924
DB 6117 CACCCTGGATGTCGACCTCTGGCAAACGGTCAATTGGAGTCCAGACAGACCCAC 6176
QY 925 TGAAGGGGCTCCGACGCTCAGGGAGGTCTCCGCCCGGACATGTGTTCAGGGACTGAG 984
DB 6177 TGAAGGGGCTCCGACGCTCAGGGAGGTCTCCGCCCGGACATGTGTTCAGGGACTGAG 6236
QY 985 CCGGGGAGCGCCCAACCCACGCTCTCTGCGGCGCACCCAGGGGAGGTAGC 1044
DB 6237 CCGGGGAGCGCCCAACCCACGCTCTCTGCGGCGCACCCAGGGGAGGTAGC 6296
QY 1045 AGCGAGGCTCTCTCTCACTCAGAGTGAACCTCAGATGTCTCACTCAGACTCAACTTGTCT 1104
DB 6297 AGCGAGGCTCTCTCTCACTCAGAGTGAACCTCAGATGTCTCACTCAGACTCAACTTGTCT 6356
QY 1105 GGGACACTGACAGAGCGTTCTGAGAGTTTTCAGCCACTTAGCGTGTGGCGTTTAAAGAT 1164
DB 6357 GGGACACTGACAGAGCGTTCTGAGAGTTTTCAGCCACTTAGCGTGTGGCGTTTAAAGAT 6416
QY 1165 CCCTCAAGTCACTGAGAAATGCGACAGATGTCAGAGAGCCTGGAGAGCTTCTGTGAG 1224
DB 6417 CCCTCAAGTCACTGAGAAATGCGACAGATGTCAGAGAGCCTGGAGAGCTTCTGTGAG 6476
QY 1225 GAATGAGGACATTTTGGGGAATTTGAGAGACAGCTTAGACATGCTGGCTGAT 1284
DB 6477 GAATGAGGACATTTTGGGGAATTTGAGAGAGAGCTTAGACATGCTGGCTGAT 6536
QY 1285 GTTTTGTGACAGTGAACCCACAGTGGAGAGAGTTTTCAGTCTGATCTGGTTCTTA 1344
DB 6537 GTTTTGTGACAGTGAACCCACAGTGGAGAGAGTTTTCAGTCTGATCTGGTTCTTA 6596
QY 1345 CAACTCAACAATAAATCTAAAAAGTTTGTGAACAAGTACTTCTTTTACATGTTA 1404
DB 6597 CAACTCAACAATAAATCTAAAAAGTTTGTGAACAAGTACTTCTTTTACATGTTA 6656
QY 1405 CATGCTCATGTTTCTGTTTCTGTTTCAATAACAAGAGTGTGTGGCTCAACAAC 1464
DB 6657 CATGCTCATGTTTCTGTTTCTGTTTCAATAACAAGAGTGTGTGTGGCTCAACAAC 6716
QY 1465 CTAAATTCATGACCCAGTGTGTCAGTCCAGCGTGGCTTACAACGATATGGGAGCCAC 1524
DB 6717 CTAAATTCATGACCCAGTGTGTCAGTCCAGCGTGGCTTACAACGATATGGGAGCCAC 6776
QY 1525 TGAAGGATGTTTCCCGCTTGTGCTTGAAGAGCAAGAGCAAGAGCGAGTGGCT 1584
DB 6777 TGAAGGATGTTTCCCGCTTGTGCTTGAAGAGCAAGAGCGAGTGGCT 6836
QY 1585 GGAAGCACCAGATCAACCCAGGCTTGTGCGGGCCAG 1624
```

```
DB 6837 GGAAGCACCAGATCAACCCAGGCTTGTGCGGGCCAG 6876

RESULT 6
US-09-896-522-3
/ Sequence 3, Application US/09896522
/ Patent No. US20020055161A1
/ GENERAL INFORMATION:
/ APPLICANT: Gluckmann, Maria A.
/ TITLE OF INVENTION: A NOVEL HUMAN URIDINE KINASE AND
/ TITLE OF INVENTION: USES THEREOF
/ FILE REFERENCE: 381552001700
/ CURRENT APPLICATION NUMBER: US/09/896,522
/ CURRENT FILING DATE: 2001-06-28
/ PRIOR APPLICATION NUMBER: 60/216,503
/ PRIOR FILING DATE: 2000-06-30
/ NUMBER OF SEQ ID NOS: 6
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 3
/ LENGTH: 834
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-896-522-3

Query Match      51.4%; Score 834; DB 9; Length 834;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 834; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 94 ATGCTTCGCGGAGGCGAAGACTGCGAGAGCCCGCGCGGAGGCGGACCGTCCGAC 153
DB 1 ATGCTTCGCGGAGGCGAAGACTGCGAGAGCCCGCGCGGAGGCGGACCGTCCGAC 60
QY 154 CAGCGGCCCTTCTCTGATAGAGGGTGAAGCGCGGCGGACCTGCGAGCGGGAAGTCAACCTGTCT 213
DB 61 CAGCGGCCCTTCTCTGATAGAGGGTGAAGCGCGGCGGACCTGCGAGCGGGAAGTCAACCTGTCT 120
QY 214 GAGAATCATGAGATGTCGTCGGAAGAGAGTGAACAGCGGAGCGGAGGAGTGTCT 273
DB 121 GAGAATCATGAGATGTCGTCGGAAGAGAGTGAACAGCGGAGCGGAGGAGTGTCT 180
QY 274 ATCTTAGGCAAGACAGTCTTCAACAGTCTTGAACGCGAGAGCAAGAGGCGCTTG 333
DB 181 ATCTTAGGCAAGACAGTCTTCAACAGTCTTGAACGCGAGAGCAAGAGGCGCTTG 240
QY 334 AAAGACATGACATTTTGAACCATCCAGATGCTTGAATATGATTTGATGACAGACT 333
DB 241 AAAGACATGACATTTTGAACCATCCAGATGCTTGAATATGATTTGATGACAGACT 300
QY 394 CTGAAGACATGCTGAGGCGAAGAGCGTGAAGTCCGACCTATGATTTTGAACAC 453
DB 301 CTGAAGACATGCTGAGGCGAAGAGCGTGAAGTCCGACCTATGATTTTGAACAC 360
QY 454 TCAAGTTTACAGAGACACAGTGTCTTACCTGCGAGAGTGTCTGTTTGAAGGCATC 513
DB 361 TCAAGTTTACAGAGACACAGTGTCTTACCTGCGAGAGTGTCTGTTTGAAGGCATC 420
QY 514 TTGCTTTTCAACCCAGAGATCCGGGACATGTTTCACTGCGCTCTTCTGTGACACC 573
DB 421 TTGCTTTTCAACCCAGAGATCCGGGACATGTTTCACTGCGCTCTTCTGTGACACC 480
QY 574 GACTCGAGTCAAGGCTGCTGAAGAGTCTCCGGGAGTGTGCGGAGGAGGAGGAGCTG 633
DB 481 GACTCGAGTCAAGGCTGCTGAAGAGTCTCCGGGAGTGTGCGGAGGAGGAGGAGCTG 540
QY 634 GAGCAGATTCGACGAGTACACCACTTCTGTAAGACCGGCTTCAAGAGTTCCTG 693
DB 541 GAGCAGATTCGACGAGTACACCACTTCTGTAAGACCGGCTTCAAGAGTTCCTG 600
QY 694 CCGACAAAGAGATGTCGATGATCATCCACGAGAGTGAACAAATATGTTGCTATC 753
DB 601 CCGACAAAGAGATGTCGATGATCATCCACGAGAGTGAACAAATATGTTGCTATC 660
QY 754 AACCTGATGTGAGACATCCAGACATTTGAAATGTGACATCTGCAAAATGACACGA 813
```

Db 661 AACCTGATCGTGCAGCATTCAGAGACATTTGATGTGACATCTGCAAAATGACACGA 720
Qy 814 GAGAGGTTCATTTGGGCGGAGCTACAGCGGACCTTTTGTGAGCCAGGGGACACCTGGG 873
Db 721 GAGAGGTTCATTTGGGCGGAGCTACAGCGGACCTTTTGTGAGCCAGGGGACACCTGGG 780
Qy 874 ATGCTGACCTCTGGCAAAAGGTGACATTTGGAGTCCAGCAGACAGCCCACTGA 927
Db 781 ATGCTGACCTCTGGCAAAAGGTGACATTTGGAGTCCAGCAGACAGCCCACTGA 834

RESULT 7

US-10-029-386-2183
; Sequence 2183, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR C
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AEMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 2183
; LENGTH: 510
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR9,1
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.5
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.8
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.6
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.7
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.5
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.7
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.3
; OTHER INFORMATION: NT HIT: g113899252, EVALUE 0.00e+00
; OTHER INFORMATION: EST HUMAN HIT: BF664526.1, EVALUE 0.00e+00
; OTHER INFORMATION: SWISSPROT HIT: P52623, EVALUE 2.00e-25
US-10-029-386-2183

Query Match 29.9%; Score 485; DB 12; Length 510;
Best Local Similarity 100.0%; Pred. No. 3.3e-250; Indels 0; Gaps 0;

Matches 485; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 745 GTTSCATCAACCTGATCGTGCAGACATTCAGAGACATTTGATGTGACATCTGCAAA 804
Db 26 GTTSCATCAACCTGATCGTGCAGACATTCAGAGACATTTGATGTGACATCTGCAAA 85
Qy 805 TGGCAGCAGAGAGGCTCAATGGCGGAGCTACAGCGGACCTTTTCTGAGCCAGGGGAC 864
Db 86 TGGCAGCAGAGAGGCTCAATGGCGGAGCTACAGCGGACCTTTTCTGAGCCAGGGGAC 145
Qy 865 CACCTTGAGATGCTGACCTCTGGCAAAAGGTGACATTTGAGAGTCCAGACAGACCCAC 924
Db 146 CACCTTGAGATGCTGACCTCTGGCAAAAGGTGACATTTGAGAGTCCAGACAGACCCAC 205
Qy 925 TGAAGGGCTGCGAGGCTCAGGGGAGGCTCCGCGCGGACATGATGTTCAGGGGAGTGGAG 984
Db 206 TGAAGGGCTGCGAGGCTCAGGGGAGGCTCCGCGCGGACATGATGTTCAGGGGAGTGGAG 265
Qy 985 CTTGGGAGCGCCACCAACCACTGCTTCTCTGCGCGGACCCAGGGGAGTGGAG 1044
Db 266 CTTGGGAGCGCCACCAACCACTGCTTCTCTGCGCGGACCCAGGGGAGTGGAG 325
Qy 1045 AGCGAGGCTTCT 1104
Db 326 AGCGAGGCTTCT 385

Qy 1105 GGGACACTGACAGAGCGTTCCTGAGGTTTTCAGGCCATTAAGGCTGTTGGGTTAAAGAT 1164
Db 386 GGGACACTGACAGAGCGTTCCTGAGGTTTTCAGGCCATTAAGGCTGTTGGGTTAAAGAT 445
Qy 1165 CCCTTAGTCTACTGAGAAAATGCGACAGAAATGTGACAGAAAGCTTGGAGGCTTCTGTAG 1224
Db 446 CCCTTAGTCTACTGAGAAAATGCGACAGAAATGTGACAGAAAGCTTGGAGGCTTCTGTAG 505
Qy 1225 GAATG 1229
Db 506 GAATG 510

RESULT 8

US-09-918-995-30379
; Sequence 30379, Application US/0918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hvaeq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-03-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 30379
; LENGTH: 472
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(472)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-30379

Query Match 24.8%; Score 403; DB 11; Length 472;
Best Local Similarity 100.0%; Pred. No. 4.5e-206; Indels 0; Gaps 0;

Matches 403; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 190 GCCAGCGGAGAGTGCACCGTGTGTGAGAAATCATGATGATGCTGTGGACAGAAAGAGGTG 249
Db 41 GCCAGCGGAGAGTGCACCGTGTGTGAGAAATCATGATGATGCTGTGGACAGAAAGAGGTG 100
Qy 250 GAAACGCGGAGCGGAGAGGTGTCTATCTGAGCCAGAGAGTCTTCAAGAGTCTGACG 309
Db 101 GAAACGCGGAGCGGAGAGGTGTCTATCTGAGCCAGAGAGTCTTCAAGAGTCTGACG 160
Qy 310 GCAGACAGAAAGGCCAAGGCTTGAAGAGCAGTACAAATTTTGAACATTCAGATGCTTT 369
Db 161 GCAGACAGAAAGGCCAAGGCTTGAAGAGCAGTACAAATTTTGAACATTCAGATGCTTT 220
Qy 370 GATATGATTTTATGACAGAGTCTGAGAAACATGCTGAGAGGCAAAAACGATGAGGTG 429
Db 221 GATATGATTTTATGACAGAGTCTGAGAAACATGCTGAGAGGCAAAAACGATGAGGTG 280
Qy 430 CCGACCTATGATTTTGTGACACATCAAGGTTACAGAGACCAAGGTTGTTACCTTGGC 489
Db 281 CCGACCTATGATTTTGTGACACATCAAGGTTACAGAGACCAAGGTTGTTACCTTGGC 340
Qy 490 GAGTGTCTCTGTTTGAAGGACATCTTGTGTCTTACAGCAGAGATCCGGGACATGTTT 549
Db 341 GAGTGTCTCTGTTTGAAGGACATCTTGTGTCTTACAGCAGAGATCCGGGACATGTTT 400
Qy 550 CACTGCGCTCTTCTGTGAGACCGAGCTCGAGAGTCTGAGGCTGT 592
Db 401 CACTGCGCTCTTCTGTGAGACCGAGCTCGAGAGTCTGAGGCTGT 443

RESULT 9

```
US-10-029-386-15883
; Sequence 15883, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR C
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 15883
; LENGTH: 187
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR9.1
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.5
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.8
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.6
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.7
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.5
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.7
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.3
; OTHER INFORMATION: EST HUMAN HIT: A1992171.1, EVALUATE 1.00e-100
; OTHER INFORMATION: SWISSPROT HIT: P52623, EVALUATE 4.00e-26
; OTHER INFORMATION: NT HIT: g14783235, EVALUATE 1.00e-100
US-10-029-386-15883
```

```
Query Match 11.4%; Score 185; DB 12; Length 187;
Best Local Similarity 100.0%; Pred. No. 1.1e-88;
Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 745 GTTGCCATCAACGTCATCGTCAGACATCCAGACATTCGTAATGGTCATCTGCAGAA 804
DB 3 GTTGCCATCAACGTCATCGTCAGACATCCAGACATTCGTAATGGTCATCTGCAGAA 62
QY 805 TGGCACCAGAGAGAGGTCATCGTCAGAGCTTACAGCGGACCTTTCTGAGCCAGGGGAC 864
DB 63 TGGCACCAGAGAGAGGTCATCGTCAGAGCTTACAGCGGACCTTTCTGAGCCAGGGGAC 122
QY 865 CACCTCGGATGCTGACCTCTGGCAACGGTCACTTTGAGTCCAGACAGACCCAC 924
DB 123 CACCTCGGATGCTGACCTCTGGCAACGGTCACTTTGAGTCCAGACAGACCCAC 182
QY 925 TGAGG 929
DB 183 TGAGG 187
```

```
RESULT 10
US-09-764-877-399
; Sequence 399, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005
; CURRENT APPLICATION NUMBER: US/09/764,877
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 399
; LENGTH: 222
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
```

```
; LOCATION: (215)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-764-877-399
```

```
Query Match 3.6%; Score 58; DB 10; Length 222;
Best Local Similarity 100.0%; Pred. No. 2.4e-20;
Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1521 CCACGTAGGAGATTTTCCCTGCTGCTGCTTAAGCAGAGACGAGCGCGGA 1578
DB 32 CCACGTAGGAGATTTTCCCTGCTGCTGCTTAAGCAGAGACGAGCGCGGA 89
```

```
RESULT 11
US-09-908-975-24159
; Sequence 24159, Application US/09908975
; Publication No. US20030165843A1
; GENERAL INFORMATION:
; APPLICANT: SHOSHAN, Avi
; APPLICANT: WASSERMAN, Alon
; APPLICANT: MINTZ, Eli
; APPLICANT: FAIGLER, Simchon
; TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICING
; FILE REFERENCE: 36688-0005
; CURRENT APPLICATION NUMBER: US/09/908,975
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: US 60/287,724
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: US 60/221,607
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 32337
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 24159
; LENGTH: 65
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-908-975-24159
```

```
Query Match 1.6%; Score 26; DB 12; Length 65;
Best Local Similarity 100.0%; Pred. No. 0.0043;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 586 AGGCTGTCTCGAAGATTTCTCCGGGA 611
DB 16 AGGCTGTCTCGAAGATTTCTCCGGGA 41
```

```
RESULT 12
US-09-925-300-220
; Sequence 220, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Ruben
; APPLICANT: Steve Ruben
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 220
; LENGTH: 1310
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-300-220
```

```
Query Match 1.4%; Score 23; DB 10; Length 1310;
```

Best Local Similarity 100.0%; Pred. No. 0.16; Indels 0; Gaps 0;
Matches 23; Conservative 0; Mismatches 0;

QY 760 ATCGTCAGCATCCAGACAT 782
Db 891 ATCGTCAGCATCCAGACAT 913

RESULT 13

US-10-098-841-53
; Sequence 53, Application US/10098841
; Publication No. US20020197679A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Xu, Chongjun
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yungqing
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhao, Qing A.
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: Wehrman, Tom
; APPLICANT: Zhang, Jie
; APPLICANT: Qian, Xiaohong B.
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20020197679A1e1. Nucleic Acids and
; FILE REFERENCE: 784CIP2
; CURRENT APPLICATION NUMBER: US/10/098,841
; PRIOR FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: 09/598,042
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 331
; SOFTWARE: pt_genes Version 1.0
; SEQ ID NO 53
; LENGTH: 1402
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (290)..(1075)
US-10-098-841-53

Query Match 1.4%; Score 23; DB 13; Length 1402;
Best Local Similarity 100.0%; Pred. No. 0.16; Indels 0; Gaps 0;

Matches 23; Conservative 0; Mismatches 0;

QY 760 ATCGTCAGCATCCAGACAT 782
Db 950 ATCGTCAGCATCCAGACAT 972

RESULT 14

US-10-066-543-1843
; Sequence 1843, Application US/10066543
; Publication No. US20030087818A1
; GENERAL INFORMATION:
; APPLICANT: Jjiang, Yugu
; APPLICANT: Pyle, Ruth A.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Indrias, Carol Yoseph
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.

; APPLICANT: Smith, Carole L.
; APPLICANT: Durham, Margarita
; APPLICANT: Scolk, John A.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.563
; CURRENT APPLICATION NUMBER: US/10/066,543
; CURRENT FILING DATE: 2002-01-31

; NUMBER OF SEQ ID NOS: 3417

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 1843

; LENGTH: 447

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-066-543-1843

Query Match 1.4%; Score 22; DB 14; Length 447;
Best Local Similarity 100.0%; Pred. No. 0.57; Indels 0; Gaps 0;
Matches 22; Conservative 0; Mismatches 0;

QY 266 AGGTGTCATCCTGAGCCAGGA 287
Db 177 AGGTGTCATCCTGAGCCAGGA 198

RESULT 15

US-09-918-995-23923
; Sequence 23923, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 23923
; LENGTH: 455
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(455)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-23923

Query Match 1.4%; Score 22; DB 11; Length 455;
Best Local Similarity 100.0%; Pred. No. 0.57; Indels 0; Gaps 0;

Matches 22; Conservative 0; Mismatches 0;

QY 266 AGGTGTCATCCTGAGCCAGGA 287
Db 191 AGGTGTCATCCTGAGCCAGGA 212

Search completed: November 25, 2003, 07:26:58
Job time: 1297.63 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 25, 2003, 02:04:12 ; Search time 113.64 Seconds
(Without alignments)
6307.673 Million cell updates/sec

Title: US-09-896-522-1

Sequence: 1 gggggggtcgcccccagaccctc.....ccagctctgctgggggacag 1624

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 569978 seqs, 220691566 residues

Word size : 15

Total number of hits satisfying chosen parameters: 934

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

Issued Patents NA: *
1: /cgn2_6/prodata/2/ina/5A_COMB.seq.*
2: /cgn2_6/prodata/2/ina/5B_COMB.seq.*
3: /cgn2_6/prodata/2/ina/6A_COMB.seq.*
4: /cgn2_6/prodata/2/ina/6B_COMB.seq.*
5: /cgn2_6/prodata/2/ina/PCUS_COMB.seq.*
6: /cgn2_6/prodata/2/ina/backfile1.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	* Query Match	Length	DB ID	Description
1	932	57.4	1288	4 US-09-620-312D-546	Sequence 546, App
2	783	48.2	834	4 US-09-536-647-1	Sequence 1, Appl
3	21	1.3	1529	3 US-08-858-876A-3	Sequence 3, Appl
4	21	1.3	1529	3 US-09-472-880-3	Sequence 3, Appl
5	21	1.3	3060	1 US-08-098-141-1	Sequence 1, Appl
6	19	1.2	1521	1 US-08-670-354-3	Sequence 3, Appl
7	19	1.2	1521	3 US-09-320-424-3	Sequence 3, Appl
8	19	1.2	1521	4 US-09-825-563-3	Sequence 3, Appl
9	19	1.2	1521	5 PCT-US96-10895-3	Sequence 1, Appl
10	19	1.2	1751	1 US-08-670-354-1	Sequence 1, Appl
11	19	1.2	1751	3 US-09-320-424-1	Sequence 1, Appl
12	19	1.2	1751	4 US-09-825-563-1	Sequence 1, Appl
13	19	1.2	1751	5 PCT-US96-10895-1	Sequence 1, Appl
14	19	1.2	1751	4 US-09-320-424-1	Sequence 1, Appl
15	19	1.2	1751	4 US-09-320-424-1	Sequence 1, Appl
16	19	1.2	1751	4 US-09-320-424-1	Sequence 1, Appl
17	19	1.2	1751	4 US-09-320-424-1	Sequence 1, Appl
18	19	1.2	1751	4 US-09-320-424-1	Sequence 1, Appl
19	19	1.2	1751	4 US-09-320-424-1	Sequence 1, Appl
20	19	1.2	1751	4 US-09-320-424-1	Sequence 1, Appl
21	19	1.2	1751	4 US-09-320-424-1	Sequence 1, Appl
22	19	1.2	1751	4 US-09-320-424-1	Sequence 1, Appl
23	19	1.2	1751	4 US-09-320-424-1	Sequence 1, Appl
24	19	1.2	1751	4 US-09-320-424-1	Sequence 1, Appl
25	19	1.2	1751	4 US-09-320-424-1	Sequence 1, Appl
26	19	1.2	1751	4 US-09-320-424-1	Sequence 1, Appl
27	19	1.2	1751	4 US-09-320-424-1	Sequence 1, Appl

28	18	1.1	4669	4 US-09-316-167-1	Sequence 1, Appl
29	18	1.1	4669	6 5206352-3	Patent No. 5206352
30	18	1.1	6505	2 US-08-793-610-5	Sequence 5, Appl
31	18	1.1	8630	4 US-09-306-417-1	Sequence 1, Appl
32	18	1.1	8630	4 US-09-306-417-2	Sequence 2, Appl
33	18	1.1	8630	4 US-08-793-610-6	Sequence 6, Appl
34	17	1.0	245	2 US-08-691-814B-76	Sequence 76, Appl
35	17	1.0	275	2 US-08-691-814B-79	Sequence 79, Appl
36	17	1.0	276	2 US-08-691-814B-78	Sequence 78, Appl
37	17	1.0	290	2 US-08-691-814B-80	Sequence 80, Appl
38	17	1.0	300	2 US-08-691-814B-75	Sequence 75, Appl
39	17	1.0	383	2 US-08-691-814B-74	Sequence 74, Appl
40	17	1.0	431	2 US-08-691-814B-84	Sequence 84, Appl
41	17	1.0	543	2 US-08-691-814B-81	Sequence 81, Appl
42	17	1.0	592	3 US-09-385-982-265	Sequence 265, Appl
43	17	1.0	600	4 US-09-252-991A-4602	Sequence 4602, Ap
44	17	1.0	672	4 US-09-252-991A-15543	Sequence 15543, A
45	17	1.0	749	1 US-08-579-667-9	Sequence 9, Appl

ALIGNMENTS

RESULT 1
US-09-620-312D-546
Sequence 546, Application US/09620312D
Patent No. 6569662
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chenghua
APPLICANT: Asundi, Vinod
APPLICANT: Zhang, Jie
APPLICANT: Ren, Feiyan
APPLICANT: Chen, Rui-hong
APPLICANT: Zhao, Qing A.
APPLICANT: Wehrman, Tom
APPLICANT: Xue, Aidong J.
APPLICANT: Yang, Yonghong
APPLICANT: Wang, Jian-Rui
APPLICANT: Zhou, Ping
APPLICANT: Ma, Yungting
APPLICANT: Wang, Dunrui
APPLICANT: Wang, Zhilui
APPLICANT: John Tillinghaast
APPLICANT: Drmanac, Radoje T.
TITLE OF INVENTION: No. 6569662el Nucleic Acids and
FILE REFERENCE: 784CIP2B
CURRENT APPLICATION NUMBER: US/09/620,312D
CURRENT FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: 09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/488,725
PRIOR FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 1105
SOFTWARE: pc_ftl_genes Version 1.0
SEQ ID NO 546
LENGTH: 1288
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (95)..(928)
US-09-620-312D-546

Query Match 57.4%; Score 932; DB 4; Length 1288;
Best Local Similarity 100.0%; Pred No. 0;
Matches 932; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 46 GGGGAAGGGGCGGGCGGGGAGCCGATGCGCGGAGCGGAGCCGAGATGCTTCGGCG 105
DB 47 GGGGAAGGGGCGGGCGGGGAGCCGATGCGCGGAGCGGAGCCGAGATGCTTCGGCG 106

QY 106 GGAAGCCGAAGCTGCGAGAGCCCGCCGCGGAGGCGGACCGTCCGCAACAGCGGCCCTTC 165
DB 107 GGAAGCCGAAGCTGCGAGAGCCCGCCGCGGAGGCGGACCGTCCGCAACAGCGGCCCTTC 166
QY 166 CTGATAGGGGGGAGCGCGGCACTGCGAGCGGAAAGTCSACCGTGTGAGAAATCATG 225
DB 167 CTGATAGGGGGGAGCGCGGCACTGCGAGCGGAAAGTCSACCGTGTGAGAAATCATG 226
QY 226 GAGTTCCTGGAGCAAGAAAGAGAGTGAACAGCGGACCGGAAAGTGTCTATCTTGAGCCAG 285
DB 227 GAGTTCCTGGAGCAAGAAAGAGAGTGAACAGCGGACCGGAAAGTGTCTATCTTGAGCCAG 286
QY 286 GAGAGTTTCTAACAAGTCTGACGCGAGAGCAAGAGCGCTTGAAAGAGCAATAC 345
DB 287 GAGAGTTTCTAACAAGTCTGACGCGAGAGCAAGAGCGCTTGAAAGAGCAATAC 346
QY 346 AATTTTGCATCCAGATGCTTTGATATGATTTGATGACAGAGACTTGAAAGATC 405
DB 347 AATTTTGCATCCAGATGCTTTGATATGATTTGATGACAGAGACTTGAAAGATC 406
QY 406 GTGAGAGGCAAAACGCTGAGAGTCCGACCTATGATTTTGTGACACTCAAGGTTTACA 465
DB 407 GTGAGAGGCAAAACGCTGAGAGTCCGACCTATGATTTTGTGACACTCAAGGTTTACA 466
QY 466 GAGACCAAGGTGTCTACCTGCGAGAGTGTCTGTTTGAAGGCACTTGTGTCTAC 525
DB 467 GAGACCAAGGTGTCTACCTGCGAGAGTGTCTGTTTGAAGGCACTTGTGTCTAC 526
QY 526 AGCAGAGAGATCCGAGACATGTTCCACTGCGCTCTTCTGTGACACCGACTCCGAGTCC 585
DB 527 AGCAGAGAGATCCGAGACATGTTCCACTGCGCTCTTCTGTGACACCGACTCCGAGTCC 586
QY 586 AGGCTGTCTCGAAGATTTCTCCGAGAGTGTCTGCGAGGAGAGCACTGAGCAATTTCTG 645
DB 587 AGGCTGTCTCGAAGATTTCTCCGAGAGTGTCTGCGAGGAGAGCACTGAGCAATTTCTG 646
QY 646 AGCAGATACACCACTTCTGTGAGCGGCTCTGAGAGTCTGCTGCGGCAAAAG 705
DB 647 AGCAGATACACCACTTCTGTGAGCGGCTCTGAGAGTCTGCTGCGGCAAAAG 706
QY 706 TATCCGATGTATCATCCACAGAGAGTGAACAATGTTGCAATCACTGATGCTG 765
DB 707 TATCCGATGTATCATCCACAGAGAGTGAACAATGTTGCAATCACTGATGCTG 766
QY 766 CAGCACTCCAGAGATTTCTGAATGTGTGACATCTGCAAAATGTGACACCGAGGAGTCCAAT 825
DB 767 CAGCACTCCAGAGATTTCTGAATGTGTGACATCTGCAAAATGTGACACCGAGGAGTCCAAT 826
QY 826 GGGGGAGCTCAAGCGGACCTTTTCTGAGCAGGGGACCACTGAGGATGCTGACCT 885
DB 827 GGGGGAGCTCAAGCGGACCTTTTCTGAGCAGGGGACCACTGAGGATGCTGACCT 886
QY 886 GCGAAACGCTCACTTTGAGTCTGACAGAGAGCCCACTGAGGAGTCTGAGCCCTGAG 945
DB 887 GCGAAACGCTCACTTTGAGTCTGACAGAGAGCCCACTGAGGAGTCTGAGCCCTGAG 946
QY 946 GCGAGGTTCCCGCCCGGCAATGTGTTCAGG 977
DB 947 GCGAGGTTCCCGCCCGGCAATGTGTTCAGG 978

RESULT 2
US-09-536-647-1
; Sequence 1, Application US/09536647
; Patent No. 6579708
; GENERAL INFORMATION:
; APPLICANT: Johnson, Randall
; APPLICANT: Ho, Yen Sen
; TITLE OF INVENTION: No. 6579708e1 Human Uridine Kinase
; FILE REFERENCE: GP50020
; CURRENT APPLICATION NUMBER: US/09/536,647
; CURRENT FILING DATE: 2000-03-27
; NUMBER OF SEQ ID NOS: 3

; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 834
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(780)
US-09-536-647-1
Query Match 48.2%; Score 783; DB 4; Length 834;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 833; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 94 ATGGCTTCGGCGGAGCGGAGCACTGCGAGAGCCCGCCGAGGCGGACCGTCCGAC 153
DB 1 ATGGCTTCGGCGGAGCGGAGCACTGCGAGAGCCCGCCGAGGCGGACCGTCCGAC 60
QY 154 CAGCGGCCCTTCTGATAGGGGTGAGCGGCGGCACTGCGAGCGGAAAGTCSACCGTGTG 213
DB 61 CAGCGGCCCTTCTGATAGGGGTGAGCGGCGGCACTGCGAGCGGAAAGTCSACCGTGTG 120
QY 214 GAGAAGATCATGAGAGTGTGCGGACAGAAAGAGTGAACAGCGGCGGAAAGTGTG 273
DB 121 GAGAAGATCATGAGAGTGTGCGGACAGAAAGAGTGAACAGCGGCGGAAAGTGTG 180
QY 274 ATCTTGAGCCAGGACAGGTTCTCAAGGTCTGACGCGAGAGCAAGAGCCCAAGGCTTG 333
DB 181 ATCTTGAGCCAGGACAGGTTCTCAAGGTCTGACGCGAGAGCAAGAGCCCAAGGCTTG 240
QY 334 AAGGACAGTACATTTTGTGACATCCAGATGCTTTGATATGATTTGATGACAGAGCT 393
DB 241 AAGGACAGTACATTTTGTGACATCCAGATGCTTTGATATGATTTGATGACAGAGCT 300
QY 394 CTGAAGACATCTGTGAGGCGCAAAACGCTGAGAGTCCGACCTATGATTTTGTGACAC 453
DB 301 CTGAAGACATCTGTGAGGCGCAAAACGCTGAGAGTCCGACCTATGATTTTGTGACAC 360
QY 454 TCAAGTTTACAGAGACACAGGTGTCTACCTGCGGAGTGTCTGTTTGAAGGCTATC 513
DB 361 TCAAGTTTACAGAGACACAGGTGTCTACCTGCGGAGTGTCTGTTTGAAGGCTATC 420
QY 514 TTGGTGTCTACAGCCAGGAGATCCGAGCAATGTTCCACTGCGCTCTTCTGTGACAC 573
DB 421 TTGGTGTCTACAGCCAGGAGATCCGAGCAATGTTCCACTGCGCTCTTCTGTGACAC 480
QY 574 GACTCCGACGTCAAGGCTGTCTGAAAGATTTCTCGGAGCGTGTGCGGAGGAGGACTG 613
DB 481 GACTCCGACGTCAAGGCTGTCTGAAAGATTTCTCGGAGCGTGTGCGGAGGAGGACTG 540
QY 613 GAGCAATTTTGAAGCAATCACTTCTGTAAGCCGCGCTTCTGAGAGTCTGCTG 693
DB 541 GAGCAATTTTGAAGCAATCACTTCTGTAAGCCGCGCTTCTGAGAGTCTGCTG 600
QY 694 CCGAAGAAAGATGATGCGGATGATCATCCAGAGAGTGAACAATATGTTGCTGATC 753
DB 601 CCGAAGAAAGATGATGCGGATGATCATCCAGAGAGTGAACAATATGTTGCTGATC 660
QY 754 AACCTGATGTGAGACATCCAGACATTTCTGAATGTGTGACATCTGCAATGTGACCA 813
DB 661 AACCTGATGTGAGACATCCAGACATTTCTGAATGTGTGACATCTGCAATGTGACCA 720
QY 814 GAGAGTTCATAGGGGGAGCTCAAGCGGACCTTTTCTGAGCAGGAGGACCACTTGG 873
DB 721 GAGAGTTCATAGGGGGAGCTCAAGCGGACCTTTTCTGAGCAGGAGGACCACTTGG 780
QY 874 ATGCTGACCTTGTGAGCAAGGTCACATTTGAGATCCAGAGAGAGCCCACTGA 927
DB 781 ATGCTGACCTTGTGAGCAAGGTCACATTTGAGATCCAGAGAGAGCCCACTGA 834

RESULT 3
US-08-858-876A-3/C

Sequence 3, Application US/08858876A
Patent No. 6022856
GENERAL INFORMATION:
APPLICANT: Daniel CAPUT
APPLICANT: Pascale CHARLON
APPLICANT: Pascual FERRARA
APPLICANT: Vica NATALIO
TITLE OF INVENTION: Type 2 Neurotensin Receptor
TITLE OF INVENTION: (hnt-R2)
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Jacobson, Price, Holman & Stern, PLLC
STREET: 400 Seventh Street
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/858,876A
FILING DATE: 19-SEP-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/FR 9723204
FILING DATE: 17-MAR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Player, William E.
REGISTRATION NUMBER: 31,049
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1529 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
ORIGINAL SOURCE:
ORGANISM: Rat
FEATURE:
NAME/KEY: CDS
LOCATION: 41..1288
US-08-858-876A-3
Query Match 1.3%; Score 21; DB 3; Length 1529;
Best Local Similarity 100.0%; Pred. No. 0.86;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1048 GAGGCTTCTCTCACTCAGAG 1068
DB 826 GAGGCTTCTCTCACTCAGAG 806
RESULT 4
US-09-472-880-3/C
Sequence 3, Application US/09472880
Patent No. 6274333
GENERAL INFORMATION:
APPLICANT: Daniel CAPUT
APPLICANT: Pascale CHARLON
APPLICANT: Pascual FERRARA
APPLICANT: Vica NATALIO
TITLE OF INVENTION: Type 2 Neurotensin Receptor
TITLE OF INVENTION: (hnt-R2)
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Jacobson, Price, Holman & Stern, PLLC
STREET: 400 Seventh Street
CITY: Washington
STATE: D.C.
COUNTRY: USA

ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/472,880
FILING DATE: 28-Dec-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/FR 9723204
FILING DATE: 17-MAR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Player, William E.
REGISTRATION NUMBER: 31,049
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1529 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
ORIGINAL SOURCE:
ORGANISM: Rat
FEATURE:
NAME/KEY: CDS
LOCATION: 41..1288
US-09-472-880-3
Query Match 1.3%; Score 21; DB 3; Length 1529;
Best Local Similarity 100.0%; Pred. No. 0.86;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1048 GAGGCTTCTCTCACTCAGAG 1068
DB 826 GAGGCTTCTCTCACTCAGAG 806
RESULT 5
US-08-098-141-1
Sequence 1, Application US/08098141
Patent No. 5441875
GENERAL INFORMATION:
APPLICANT: Hediger Ph.D., Matthias A.
TITLE OF INVENTION: Urea Transporter Polypeptide
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Choate, Hall & Stewart
STREET: Exchange Place, 53 State Street
CITY: Boston
STATE: MA
COUNTRY: U.S.A.
ZIP: 02119
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/098,141
FILING DATE: 19930723
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Kaplan Esq., Warren A.
REGISTRATION NUMBER: 34,199
REFERENCE/DOCKET NUMBER: 092662-001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617:227-5020
TELEFAX: 617:227-7566
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:

LENGTH: 3060 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-098-141-1

Query Match 1.3%; Score 21; DB 1; Length 3060;
Best Local Similarity 100.0%; Pred. No. 0.86;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 272 TCATCTGAGCCAGACAGCT 292
DB 1344 TCATCTGAGCCAGACAGCT 1364

RESULT 6
US-08-670-354-3/c
Sequence 3, Application US/08670354
Patent No. 5763223
GENERAL INFORMATION:
APPLICANT: Steven R. Wiley and
APPLICANT: Raymond G. Goodwin.
TITLE OF INVENTION: Cytokine That Induces Apoptosis
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kathryn A. Anderson, Immunex Corporation
STREET: 51 University Street
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Apple 7.5.2
SOFTWARE: Microsoft Word, Version 6.0.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/670,354
FILING DATE: 25-JUN-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/496,632
FILING DATE: 29-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/548,368
FILING DATE: 01-NOV-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Anderson, Kathryn A.
REGISTRATION NUMBER: 32,172
REFERENCE/DOCKET NUMBER: 2835-B
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 587-0430
TELEFAX: (206) 233-0644
TELEX: 756822
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1521 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
IMMEDIATE SOURCE:
CLONE: HUAIC-dv
FEATURE:
NAME/KEY: CDS
LOCATION: 78..383

US-08-670-354-3

Query Match 1.2%; Score 19; DB 1; Length 1521;
Best Local Similarity 100.0%; Pred. No. 8.8;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1415 TGTTTCGTGTTTCTGTTT 1433
DB 891 TGTTTCGTGTTTCTGTTT 873

RESULT 7
US-09-320-424-3/c
Sequence 3, Application US/09320424
Patent No. 6284236
GENERAL INFORMATION:
APPLICANT: Wiley, Steven R.
APPLICANT: Goodwin, Raymond G.
TITLE OF INVENTION: Cytokine that Induces Apoptosis
FILE REFERENCE: 2835-E
CURRENT APPLICATION NUMBER: US/09/320,424
CURRENT FILING DATE: 1999-05-26
EARLIER APPLICATION NUMBER: 09/190,046
EARLIER FILING DATE: 1998-11-10
EARLIER APPLICATION NUMBER: 09/048,641
EARLIER FILING DATE: 1998-03-26
EARLIER APPLICATION NUMBER: 08/670,354
EARLIER FILING DATE: 1996-06-25
EARLIER APPLICATION NUMBER: 08/548,368
EARLIER FILING DATE: 1995-11-01
EARLIER APPLICATION NUMBER: 08/496,632
EARLIER FILING DATE: 1995-06-29
NUMBER OF SEQ ID NOS: 25
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 3
LENGTH: 1521
TYPE: DNA
ORGANISM: human
FEATURE:
NAME/KEY: CDS
LOCATION: (78)..(383)
US-09-320-424-3

Query Match 1.2%; Score 19; DB 3; Length 1521;
Best Local Similarity 100.0%; Pred. No. 8.8;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1415 TGTTTCGTGTTTCTGTTT 1433
DB 891 TGTTTCGTGTTTCTGTTT 873

RESULT 8
US-09-825-563-3/c
Sequence 3, Application US/09825563
Patent No. 6521228
GENERAL INFORMATION:
APPLICANT: Wiley, Steven R.
APPLICANT: Goodwin, Raymond G.
TITLE OF INVENTION: Cytokine that Induces Apoptosis
FILE REFERENCE: 2835-E
CURRENT APPLICATION NUMBER: US/09/825,563
CURRENT FILING DATE: 2001-04-02
PRIOR APPLICATION NUMBER: 09/320,424
PRIOR FILING DATE: 1999-05-26
PRIOR APPLICATION NUMBER: 09/190,046
PRIOR FILING DATE: 1998-11-10
PRIOR APPLICATION NUMBER: 09/048,641
PRIOR FILING DATE: 1998-03-26
PRIOR APPLICATION NUMBER: 08/670,354
PRIOR FILING DATE: 1996-06-25
PRIOR APPLICATION NUMBER: 08/548,368
PRIOR FILING DATE: 1995-11-01

PRIOR APPLICATION NUMBER: 08/496,632
PRIOR FILING DATE: 1995-06-29
NUMBER OF SEQ ID NOS: 25
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 3
LENGTH: 1521
TYPE: DNA
ORGANISM: human
FEATURE:
NAME/KEY: CDS
LOCATION: (78) .. (383)
US-09-825-563-3

Query Match 1.2%; Score 19; DB 4; Length 1521;
Best Local Similarity 100.0%; Pred. No. 8.8;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1415 TGTTCCTGTTTCGTTT 1433
DB 891 TGTTCCTGTTTCGTTT 873

RESULT 9
PCT-US96-10895-3/c
Sequence 3, Application PC/TUS9610895
GENERAL INFORMATION:
APPLICANT: Immunex Corporation.
TITLE OF INVENTION: Cytokine That Induces Apoptosis
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kathryn A. Anderson, Immunex Corporation
STREET: 51 University Street
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Apple 7.5.2
SOFTWARE: Microsoft Word, Version 6.0.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/10895
FILING DATE: 25-JUN-1996
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/496,632
FILING DATE: 29-JUN-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/548,368
FILING DATE: 01-NOV-1995
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Anderson, Kathryn A.
REGISTRATION NUMBER: 32,172
REFERENCE/DOCKET NUMBER: 2835-WO
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 587-0430
TELEFAX: (206) 233-0644
TELEX: 756822
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1521 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
IMMEDIATE SOURCE:
CLONE: HuAic-dv
FEATURE:

NAME/KEY: CDS
LOCATION: 78..383
PCT-US96-10895-3

Query Match 1.2%; Score 19; DB 5; Length 1521;
Best Local Similarity 100.0%; Pred. No. 8.8;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1415 TGTTCCTGTTTCGTTT 1433
DB 891 TGTTCCTGTTTCGTTT 873

RESULT 10
US-08-670-354-1/c
Sequence 1, Application US/08670354
Patent No. 5763223
GENERAL INFORMATION:
APPLICANT: Steven R. Wiley and
APPLICANT: Raymond G. Goodwin
TITLE OF INVENTION: Cytokine That Induces Apoptosis
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kathryn A. Anderson, Immunex Corporation
STREET: 51 University Street
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Apple 7.5.2
SOFTWARE: Microsoft Word, Version 6.0.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/670,354
FILING DATE: 25-JUN-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/496,632
FILING DATE: 29-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/548,368
FILING DATE: 01-NOV-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Anderson, Kathryn A.
REGISTRATION NUMBER: 32,172
REFERENCE/DOCKET NUMBER: 2835-B
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 587-0430
TELEFAX: (206) 233-0644
TELEX: 756822
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1751 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
IMMEDIATE SOURCE:
CLONE: huAic
FEATURE:
NAME/KEY: CDS
LOCATION: 88..933
US-08-670-354-1

Query Match 1.2%; Score 19; DB 1; Length 1751;
Best Local Similarity 100.0%; Pred. No. 8.8;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1415 TGTTCGTCTCTGTT 1433
| | | | | | | | | |
Db 1049 TGTTCGTCTCTGTT 1031

RESULT 11

US-09-320-424-1/c
; Sequence 1, Application US/09320424
; Patent No. 6284236
; GENERAL INFORMATION:
; APPLICANT: Wiley, Steven R.
; APPLICANT: Goodwin, Raymond G.
; TITLE OF INVENTION: Cytokine that Induces Apoptosis
; FILE REFERENCE: 2835-E
; CURRENT APPLICATION NUMBER: US/09/320,424
; CURRENT FILING DATE: 1999-05-26
; EARLIER APPLICATION NUMBER: 09/190,046
; EARLIER FILING DATE: 1998-11-10
; EARLIER APPLICATION NUMBER: 09/048,641
; EARLIER FILING DATE: 1998-03-26
; EARLIER APPLICATION NUMBER: 08/670,354
; EARLIER FILING DATE: 1996-06-25
; EARLIER APPLICATION NUMBER: 08/548,368
; EARLIER FILING DATE: 1995-11-01
; EARLIER APPLICATION NUMBER: 08/496,632
; EARLIER FILING DATE: 1995-06-29
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1751
; TYPE: DNA
; ORGANISM: human
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (88)..(933)
US-09-320-424-1

Query Match 1.2%; Score 19; DB 3; Length 1751;
Best Local Similarity 100.0%; Pred. No. 8.8;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1415 TGTTCGTCTCTGTT 1433
| | | | | | | | | |
Db 1049 TGTTCGTCTCTGTT 1031

RESULT 12

US-09-825-563-1/c
; Sequence 1, Application US/09825563
; Patent No. 6521228
; GENERAL INFORMATION:
; APPLICANT: Wiley, Steven R.
; APPLICANT: Goodwin, Raymond G.
; TITLE OF INVENTION: Cytokine that Induces Apoptosis
; FILE REFERENCE: 2835-E
; CURRENT APPLICATION NUMBER: US/09/825,563
; CURRENT FILING DATE: 2001-04-02
; PRIOR APPLICATION NUMBER: 09/320,424
; PRIOR FILING DATE: 1999-05-26
; PRIOR APPLICATION NUMBER: 09/190,046
; PRIOR FILING DATE: 1998-11-10
; PRIOR APPLICATION NUMBER: 09/048,641
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 08/670,354
; PRIOR FILING DATE: 1996-06-25
; PRIOR APPLICATION NUMBER: 08/548,368
; PRIOR FILING DATE: 1995-11-01
; PRIOR APPLICATION NUMBER: 08/496,632
; PRIOR FILING DATE: 1995-06-29
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 1

; LENGTH: 1751
; TYPE: DNA
; ORGANISM: human
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (88)..(933)
US-09-825-563-1

Query Match 1.2%; Score 19; DB 4; Length 1751;
Best Local Similarity 100.0%; Pred. No. 8.8;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1415 TGTTCGTCTCTGTT 1433
| | | | | | | | | |
Db 1049 TGTTCGTCTCTGTT 1031

RESULT 13

PCT-US96-10895-1/c
; Sequence 1, Application PC/TUS9610895
; GENERAL INFORMATION:
; APPLICANT: Immunex Corporation.
; TITLE OF INVENTION: Cytokine That Induces Apoptosis
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kathryn A. Anderson, Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: Apple Macintosh
; SOFTWARE: Microsoft Word, Version 6.0.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/10895
; FILING DATE: 25-JUN-1996
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/496,632
; FILING DATE: 29-JUN-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/548,368
; FILING DATE: 01-NOV-1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Anderson, Kathryn A.
; REGISTRATION NUMBER: 32,172
; REFERENCE/DOCKET NUMBER: 2835-WO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 587-0430
; TELEFAX: (206) 233-0644
; TELEX: 756822
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1751 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; IMMEDIATE SOURCE:
; CLONE: huAIC
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 88..933
; PCT-US96-10895-1

Query Match 1.2%; Score 19; DB 5; Length 1751;

Best Local Similarity 100.0%; Pred. No. 8.8;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1415 TGTTCCTGTTCTGTTT 1433
|||||
Db 1049 TGTTCCTGTTCTGTTT 1031

RESULT 14

US-09-333-593A-5/c
; Sequence 5, Application US/0933593A
; Patent No. 6313269
; GENERAL INFORMATION:
; APPLICANT: DEEN, KEITH C.
; APPLICANT: MARSHALL, LISA A.
; APPLICANT: KOSHAK, AMY K.
; APPLICANT: TAN, KONG B.
; APPLICANT: TRUENH, ALEMSEGED
; TITLE OF INVENTION: TUMOR NECROSIS FACTOR RELATED RECEPTOR,
; TITLE OF INVENTION: TR6
; FILE REFERENCE: GH-50008-2
; CURRENT APPLICATION NUMBER: US/09/333,593A
; CURRENT FILING DATE: 1999-06-15
; PRIOR APPLICATION NUMBER: 08/316,625
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 08/853,684
; PRIOR FILING DATE: 1997-05-09
; PRIOR APPLICATION NUMBER: 60/041,230
; PRIOR FILING DATE: 1997-03-14
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 1769
; TYPE: DNA
; ORGANISM: HOMO SAPIENS
US-09-333-593A-5

Query Match

Best Local Similarity 1.2%; Score 19; DB 4; Length 1769;
Best Local Similarity 100.0%; Pred. No. 8.8;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1415 TGTTCCTGTTCTGTTT 1433
|||||
Db 1049 TGTTCCTGTTCTGTTT 1031

RESULT 15

US-09-505-250-1/c
; Sequence 1, Application US/09505250A
; Patent No. 6329148
; GENERAL INFORMATION:
; APPLICANT: Rosen, Glenn
; APPLICANT: Kao, Peter
; TITLE OF INVENTION: Synergistic Anti-Cancer Therapy with
; TITLE OF INVENTION: Triptolides and Death Domain Ligands
; FILE REFERENCE: SUN-109PRV2
; CURRENT APPLICATION NUMBER: US/09/505,250A
; CURRENT FILING DATE: 2000-02-15
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 1769
; TYPE: DNA
; ORGANISM: H. sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (88)...(933)
; OTHER INFORMATION: Human TRAIL Coding Sequence
US-09-505-250-1

Query Match 1.2%; Score 19; DB 4; Length 1769;
Best Local Similarity 100.0%; Pred. No. 8.8;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1415 TGTTCCTGTTCTGTTT 1433
|||||
Db 1049 TGTTCCTGTTCTGTTT 1031

Search completed: November 25, 2003, 06:54:13
Job time : 119.64 secs

This Page Blank (uspto)

XX 10-JUN-2002.
PD 28-JUN-2001; 2001WO-US21063.
XX 30-JUN-2000; 2000US-216503P.
XX (MILL-) MILLENNIUM PHARM INC.
PA Gluckemann MA;
PI WPI; 2002-140091/18.
DR P-PSDB; AAE16592.
XX
XX New isolated human uridine kinase family polypeptide 57658, useful for
PT treating hematopoietic neoplastic disorders and disorders of neurons,
PT heart and blood vessels
XX
XX Claim 1; Fig 1a; 103pp; English.
XX
CC The patent discloses human uridine kinase-like polypeptides, designated
CC 57658 and polynucleotides encoding such proteins. 57658 DNAs and proteins
CC are useful for developing diagnostic and therapeutic agents for 57658-
CC mediated or related disorders such as hematopoietic neoplastic disorders
CC (e.g. leukemia), hematopoietic disorders (e.g. psoriasis, dermatitis,
CC diabetes mellitus, rheumatoid arthritis, disorders of bone (e.g. rickets,
CC osteoporosis), heart (e.g. sarcoma, myocardial infarction), blood vessels
CC (e.g. hypertension, atherosclerosis) and neurons (e.g. Huntington's
CC disease, Alzheimer's disease, Parkinson's disease). 57658 sequences or
CC their antibodies are useful in screening assays, detection assays (e.g.
CC forensic biology) and predictive medicine (e.g. diagnostic assays).
CC prognostic assays, and monitoring clinical trials and pharmacogenomics).
CC They are useful as reagents for diagnosing and treating 57658-mediated
CC disorders. 57658 DNAs are useful in gene therapy, in chromosome mapping,
CC to identify an individual from a minute biological sample (tissue typing)
CC and to aid in forensic identification of the biological sample. The
CC present sequence is a DNA encoding human 57658 protein.
XX
SQ Sequence 1624 BP; 354 A; 427 C; 506 G; 337 T; 0 other:
Query Match 100.0%; Score 1624; DB 24; Length 1624;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1624; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GTGGGCTGCTCCGACCTCGGCGCTGGGCGGCGCGCGCGCGGGAAGGGCGCGC 60
DB 1 GTGGGCTGCTCCGACCTCGGCGCTGGGCGGCGCGCGCGCGGGAAGGGCGCGC 60
QY GCGGGACCCGATGCGCGGAGCGGAGCGGAGATGGCTTGCGGCGGAGCGGAGATGCG 120
DB 61 GCGGGACCCGATGCGCGGAGCGGAGCGGAGATGGCTTGCGGCGGAGCGGAGATGCG 120
QY 121 GAGAGCCCGCGCGCGGAGCGGAGCGGAGATGGCTTGCGGAGCGGAGATGCG 180
DB 121 GAGAGCCCGCGCGCGGAGCGGAGCGGAGATGGCTTGCGGAGCGGAGATGCG 180
QY 181 GCGCGCACTGCGGAGCGGAGCGGAGATGGCTTGCGGAGCGGAGATGCG 240
DB 181 GCGCGCACTGCGGAGCGGAGCGGAGATGGCTTGCGGAGCGGAGATGCG 240
QY 241 AAGAGGAGGAGGAGCGGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 300
DB 241 AAGAGGAGGAGGAGCGGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 300
QY 301 GTCTGACGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 360
DB 301 GTCTGACGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 360
QY 361 GATGCTTTGATTAATGATTTGATGACAGGAGCTTGAAAGAACATGCTGGAGGAGGAG 420
DB 361 GATGCTTTGATTAATGATTTGATGACAGGAGCTTGAAAGAACATGCTGGAGGAGGAG 420
QY 421 GTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 480

DB 421 GTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 480
QY 481 TACCTGCGGAGCGTGCTGCTTTGAGGAGATCTTGCTTCAACGACGAGATCCGG 540
DB 481 TACCTGCGGAGCGTGCTGCTTTGAGGAGATCTTGCTTCAACGAGGAGATCCGG 540
QY 541 GACATGTTCACTGCGGAGCGTGCTGCTTTGAGGAGATCTTGCTTCAACGAGGAG 600
DB 541 GACATGTTCACTGCGGAGCGTGCTGCTTTGAGGAGATCTTGCTTCAACGAGGAG 600
QY 601 GTTCTCGGAGCGTGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 660
DB 601 GTTCTCGGAGCGTGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 660
QY 661 TTGTTGAGGCGGCGCTTTCGAGAGGTTCTGCTGCGGACAAAGAAATGCTCCATGATC 720
DB 661 TTGTTGAGGCGGCGCTTTCGAGAGGTTCTGCTGCGGACAAAGAAATGCTCCATGATC 720
QY 721 ATCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 780
DB 721 ATCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 780
QY 781 ATTCTGAATGTTGATCATCTGCAAAATGCGACCGAGAGGTTCCAAATGCGGAGCTCAAG 840
DB 781 ATTCTGAATGTTGATCATCTGCAAAATGCGACCGAGAGGTTCCAAATGCGGAGCTCAAG 840
QY 841 CGGACCTTTTCTGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 900
DB 841 CGGACCTTTTCTGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 900
QY 901 TTGAGTTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 960
DB 901 TTGAGTTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 960
QY 961 CGGACATGTTGTTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1020
DB 961 CGGACATGTTGTTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1020
QY 1021 GCGGACCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1080
DB 1021 GCGGACCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1080
QY 1081 TGTGTACTGAGCTCAACTGCTGGAGCACTGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1140
DB 1081 TGTGTACTGAGCTCAACTGCTGGAGCACTGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1140
QY 1141 TTAGGCTGCTGCGGTTTAAAGATCCCTCTAGGTCACTGAGAAATGCCAGAAATGTGCA 1200
DB 1141 TTAGGCTGCTGCGGTTTAAAGATCCCTCTAGGTCACTGAGAAATGCCAGAAATGTGCA 1200
QY 1201 GGAAGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1260
DB 1201 GGAAGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1260
QY 1261 AGCTTGAACACTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1320
DB 1261 AGCTTGAACACTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1320
QY 1321 TTTTCAAGTGTGATCTGCTTTTCAACACTCACTCAACATTAATCAAAAGTTTGTGAA 1380
DB 1321 TTTTCAAGTGTGATCTGCTTTTCAACACTCACTCAACATTAATCAAAAGTTTGTGAA 1380
QY 1381 AGTACTTCTCTTTTAAATGATGATGCTGCTGATGCTTCTGTTTCTGTTTCAATACA 1440
DB 1381 AGTACTTCTCTTTTAAATGATGATGCTGCTGATGCTTCTGTTTCTGTTTCAATACA 1440
QY 1441 CAAGGCTGTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1500
DB 1441 CAAGGCTGTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1500
QY 1501 GCTTACAGGATTAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1560

DB 1501 GCCTACCGGATATGAGGAGCCACTAGAGGATGTTTCCCTGCTTGCTTAAG 1560
QY 1561 GCAGAGAGCGAGCGCGATGCCCTGGAAGACCCAGATACACCCAGGTTGTGGGG 1620
DB 1561 GCAGAGAGCGAGCGCGATGCCCTGGAAGACCCAGATACACCCAGGTTGTGGGG 1620
QY 1621 CCAG 1624
DB 1621 CCAG 1624

RESULT 2
AA160444
ID AA160444 standard; cDNA; 2152 BP.
AC AA160444;
XX
XX 22-OCT-2001 (first entry)
DE Human polynucleotide SEQ ID NO 4433.
XX
XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukemia; ss.
XX
XX Homo sapiens.
OS
XX
XX W020015312-A1.
PN
PD 26-JUL-2001.
XX
XX 26-DEC-2000; 2000MO-US34263.
PF
XX
XX 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
XX
XX (HYSE-) HYSEQ INC.
PA
PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
XX WPI; 2001-442253/47.
DR P-PSDB; AAM41288.
XX
XX Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
XX
XX Claim 1; SEQ ID NO 4433; 10078BP; English.
XX
XX The invention relates to human nucleic acids (AA157798-AA161369) and
CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
XX immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukemias and
CC C.N.S disorders.

CC Note: The sequence data for this patent did not form part of the printed
CC specification.

XX Sequence 2152 BP; 468 A; 581 C; 660 G; 442 T; 1 other;

QY Query Match 94.1%; Score 1528; DB 22; Length 2152;

DB Best Local Similarity 99.9%; Pred. No. 0;

QY Matches 1578; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 46 GGGGAAAGGGCGCGCGCGGAGCCCGATGCGCGGAGCCGAGCCGAGATGCTTGGCG 105
DB 11 GGGGAAAGGGCGCGCGCGGAGCCCGATGCGCGGAGCCGAGCCGAGATGCTTGGCG 70
QY 106 GGAAGCGGAAGACTGTGAGAGCCCGCGCGGAGCCCGATGCGCGGAGCCGAGATGCTT 165
DB 71 GGAAGCGGAAGACTGTGAGAGCCCGCGCGGAGCCCGATGCGCGGAGCCGAGATGCTT 130
QY 166 CTGATAGGGGTGAGGCGCGGAGCTGCGAGCGGAGAGTCGACCGTGTGTGAGAGATCATG 225
DB 131 CTGATAGGGGTGAGGCGCGGAGCTGCGAGCGGAGAGTCGACCGTGTGTGAGAGATCATG 190
QY 226 GAGTTGCTGGGACAGACAGAGGTGAAACGCGGACCGGAGAGTGTCTTGAAGCCAG 285
DB 191 GAGTTGCTGGGACAGACAGAGGTGAAACGCGGACCGGAGAGTGTCTTGAAGCCAG 250
QY 286 GACAGGTTCTAAGGTCCTGACGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 345
DB 251 GACAGGTTCTAAGGTCCTGACGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 310
QY 346 AATTTTGACATCCAGATGCGCTTTGATATGATTTGATGACAGAGACTCTGAAGACATC 405
DB 311 AATTTTGACATCCAGATGCGCTTTGATATGATTTGATGACAGAGACTCTGAAGACATC 370
QY 406 GTGAGAGGCAAAACGCTGAGGAGTCCGACCTATGATTTGTGACACACTCAAGTTACCA 465
DB 371 GTGAGAGGCAAAACGCTGAGGAGTCCGACCTATGATTTGTGACACACTCAAGTTACCA 430
QY 466 GAGACACAGGTCGTCTACCCCTGCGGACGCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 525
DB 431 GAGACACAGGTCGTCTACCCCTGCGGACGCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 490
QY 526 AGCCAGAGATCCGGGAGCATGTTCCACTGCGCTCTTGTGTGAGACACGACTCCGACGTC 585
DB 491 AGCCAGAGATCCGGGAGCATGTTCCACTGCGCTCTTGTGTGAGACACGACTCCGACGTC 550
QY 586 AGGCTGTCTGAAGAGTTCTCCGGAAGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 645
DB 551 AGGCTGTCTGAAGAGTTCTCCGGAAGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 610
QY 646 AGCCAGTACACACCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 705
DB 611 AGCCAGTACACACCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 670
QY 706 TATGCCGATGATCATCCAGAGAGTGACAAATATGTTGTCATCAACTGATCTG 765
DB 671 TATGCCGATGATCATCCAGAGAGTGACAAATATGTTGTCATCAACTGATCTG 730
QY 766 GAGCAGATCCAGAGCATTTGTAATGAGTGCATCTGCAATATGCAACCGAGAGGAGGAG 825
DB 731 GAGCAGATCCAGAGCATTTGTAATGAGTGCATCTGCAATATGCAACCGAGAGGAGGAG 790
QY 826 GGGCGGAGTACAAACCGGACCTTTCTGAGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 885
DB 791 GGGCGGAGTACAAACCGGACCTTTCTGAGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 850
QY 886 GGCAGAGGTCATCATTTGAGTGCAGAGCAGACCCACTGAGGAGGAGGAGGAGGAGGAGGAG 945
DB 851 GGCAGAGGTCATCATTTGAGTGCAGAGCAGACCCACTGAGGAGGAGGAGGAGGAGGAGGAG 910
QY 946 GGCAGAGTCTCCGCGCGGAGTGTGTTGAGGAGTGAAGCTGGGAGGAGGAGGAGGAGGAG 1005
DB 911 GGCAGAGTCTCCGCGCGGAGTGTGTTGAGGAGTGAAGCTGGGAGGAGGAGGAGGAGGAGGAG 970

```
QY 1006 CCACTGCTTCTCTCGGCGGACCCCGAGGGAGTGTAGCGAGGCGCTTCTCACTCAG 1065
DB 971 CCACTGCTTCTCTCGGCGGACCCCGAGGGAGTGTAGCGAGGCGCTTCTCACTCAG 1030
QY 1066 GAGTGAAGAACTCAGATGTCTACTCAGACTCACTTCCTGGGACACAGAGCGTTCT 1125
DB 1031 GAGTGAAGAACTCAGATGTCTACTCAGACTCACTTCCTGGGACACAGAGCGTTCT 1090
QY 1126 GAGGTTTTCAGCCACTTATGCTGTTGGGTTTAAAGATCCCTCTAGCTCATGGAAT 1185
DB 1091 GAGGTTTTCAGCCACTTATGCTGTTGGGTTTAAAGATCCCTCTAGCTCATGGAAT 1150
QY 1186 GCCACAGAAATGTGAGGAGCGCTTCTGTGAGGAATGTGAGGACATTAATTGG 1245
DB 1151 GCCACAGAAATGTGAGGAGCGCTTCTGTGAGGAATGTGAGGACATTAATTGG 1210
QY 1246 GGAATTTGAGGAGACAGCTTACACACTGCTGCTGATGTTTGTGACAGTGAACCA 1305
DB 1211 GGAATTTGAGGAGACAGCTTACACACTGCTGCTGATGTTTGTGACAGTGAACCA 1270
QY 1306 CAGTGGAGAGAGTTTTCAGTCTGATCTGTTTTCACACTCACTCACTCACTCACTCA 1365
DB 1271 CAGTGGAGAGAGTTTTCAGTCTGATCTGTTTTCACACTCACTCACTCACTCACTCA 1330
QY 1366 AAAGTTTGTGAGAGAGTCTTCTTTCATGTTACATGTCCTCATGTTTCTGTT 1425
DB 1331 AAAGTTTGTGAGAGAGTCTTCTTTCATGTTACATGTCCTCATGTTTCTGTT 1390
QY 1426 TTCTGTTTCATACACAGAGCTGTTGTGCTTCAAACTAATTTGATGACCCAGTGT 1485
DB 1391 TTCTGTTTCATACACAGAGCTGTTGTGCTTCAAACTAATTTGATGACCCAGTGT 1450
QY 1486 TTGAGTTCAGGCTGCTGCTTACACAGGATGTGGGAGCCACTGAGGAGTGTTCCTCC 1545
DB 1451 TTGAGTTCAGGCTGCTGCTTACACAGGATGTGGGAGCCACTGAGGAGTGTTCCTCC 1510
QY 1546 GCTTGTGCTTAAAGGAGAGAGAGGAGGAGGATGCCCTGGAACACCCAGCATCACACC 1605
DB 1511 GCTTGTGCTTAAAGGAGAGAGAGGAGGAGGATGCCCTGGAACACCCAGCATCACACC 1570
QY 1606 CAGGCTTGTGCGGGGCCAG 1624
DB 1571 CAGGCTTGTGCGGGGCCAG 1589

RESULT 3
AAH15621
ID AAH15621 standard; cDNA; 2160 BP.
XX
AC AAH15621;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human cDNA sequence SEQ ID NO:13951.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-0116126.
XX
PR 29-JUL-1999; 99JP-0248036.
PR 27-AUG-1999; 99JP-0300253.
PR 11-JAN-2000; 2000JP-0118776.
PR 02-MAY-2000; 2000JP-0183767.
PR 09-JUN-2000; 2000JP-0241899.
XX
PA (HELI-) HELIX RES INST.
XX
```

```
PI Oka T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX WPI; 2001-318749/34.
DR
XX Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs -
XX
XX Claim 8; SEQ ID 13951; 2537bp + CD ROM; English.
XX
XX The present invention describes primer sets for synthesizing 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
XX of the present invention.
XX
SQ Sequence 2160 BP; 457 A; 591 C; 671 G; 441 T; 0 other;
XX
XX
XX Query Match 80.3%; Score 1304; DB 22; Length 2160;
XX Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Gaps 0;
XX Matches 1304; Conservative 0; Indels 0; Gaps 0;
XX
QY 46 GGGGAAAGGGGCGGCGCGGAGCCCGATGCGCGGAGCGGAGCCGAGATGCTTGGCG 105
DB 47 GGGGAAAGGGGCGGCGCGGAGCCCGATGCGCGGAGCGGAGCCGAGATGCTTGGCG 106
QY 106 GAGGCGAAGACTGCGAGAGCCCGCGCGGAGCCGAGCCGATCCGACACAGCGGCCCTTC 165
DB 107 GAGGCGAAGACTGCGAGAGCCCGCGCGGAGCCCGAGCCGAGCCGATCCGACACAGCGGCCCTTC 166
QY 166 CTGATAGGGGTGAGCGGCGGCACTGCGAGGGGAGTGCACCGTGTGAGGAAGATCATG 225
DB 167 CTGATAGGGGTGAGCGGCGGCACTGCGAGGGGAGTGCACCGTGTGAGGAAGATCATG 226
QY 226 GAGTTGCTGGAGCAGAACGAGTGTGAACAGCGGACGGAAGGTGTCTCTGAGCCAG 285
DB 227 GAGTTGCTGGAGCAGAACGAGTGTGAACAGCGGACGGAAGGTGTCTCTGAGCCAG 286
QY 286 GACAGGTTCTACAAAGTCTGACGCGACGACAGAAAGGCCCTTGAAGAAGACGTAC 345
DB 287 GACAGGTTCTACAAAGTCTGACGCGACGACAGAAAGGCCCTTGAAGAAGACGTAC 346
QY 346 AATTTTGACCATCAGATGCTTGAATGATTTATGATGACAGGACTGGAAGACATC 405
DB 347 AATTTTGACCATCAGATGCTTGAATGATTTATGATGACAGGACTGGAAGACATC 406
QY 406 GTGAGGCGCAAAACGCTGAGGTGCGGACCTATGATTTTGTGACACACTCAAGTTACCA 465
DB 407 GTGAGGCGCAAAACGCTGAGGTGCGGACCTATGATTTTGTGACACACTCAAGTTACCA 466
QY 466 GAGACCAAGGTGTCTACCTCTGCGAGCGTGTCTGTTTGAAGGCATCTTGTTCTAC 525
DB 467 GAGACCAAGGTGTCTACCTCTGCGAGCGTGTCTGTTTGAAGGCATCTTGTTCTAC 526
```

```

QY 526 AGCCAGAGATCCGGACATGTTCCACCTGCGCTCTTTCGAGACACCGACTCCGAGCTC 585
  |||
Db 527 AGCCAGAGATCCGGACATGTTCCACCTGCGCTCTTTCGAGACACCGACTCCGAGCTC 586
QY 586 AGGCTGCTCGAAGAGTTCTCCGGACGTGCGCGAGAGAGGACCTGAGACAGATTCTG 645
  |||
Db 587 AGGCTGCTCGAAGAGTTCTCCGGACGTGCGCGAGAGAGGACCTGAGACAGATTCTG 646
QY 646 ACGGAGTACACACCTCTGTGAAACCGGCTTGCAGAGAGTTCTGCTGCGACAAAGAG 705
  |||
Db 647 ACGGAGTACACACCTCTGTGAAACCGGCTTGCAGAGAGTTCTGCTGCGACAAAGAG 706
QY 706 TATCCGATGTGATCATCCACGAGAGTGGACAAATATGTTGCATCAACTGATGCTG 765
  |||
Db 707 TATCCGATGTGATCATCCACGAGAGTGGACAAATATGTTGCATCAACTGATGCTG 766
QY 766 CAGCAGATCCAGGACATTTCTGAAATGTGACATCTGCAATGGACCGAGAGGCTCCAT 825
  |||
Db 767 CAGCAGATCCAGGACATTTCTGAAATGTGACATCTGCAATGGACCGAGAGGCTCCAT 826
QY 826 GGGGCGAGCTACAAACCGGACCTTTTCTGAGCCAGGGGACACCTCGGATGCTGACTCT 885
  |||
Db 827 GGGGCGAGCTACAAACCGGACCTTTTCTGAGCCAGGGGACACCTCGGATGCTGACTCT 886
QY 886 GGCAGAACGTCACATTTGAGTCCAGGACGAGACCCGACCTGAGGGGCTGCCGAGCTCAG 945
  |||
Db 887 GGCAGAACGTCACATTTGAGTCCAGGACGAGACCCGACCTGAGGGGCTGCCGAGCTCAG 946
QY 946 GGCAGGTCTCCCGCCGCGCATGTGTGTCAGAGGACTGAGCCTGGGAGCGCCACCCAC 1005
  |||
Db 947 GGCAGGTCTCCCGCCGCGCATGTGTGTCAGAGGACTGAGCCTGGGAGCGCCACCCAC 1006
QY 1006 CCAATGCTTCTCTCTGCGCGCACCCGAGGGAGTGTTCAGCGAGGCTTCTCACTCAG 1065
  |||
Db 1007 CCAATGCTTCTCTCTGCGCGCACCCGAGGGAGTGTTCAGCGAGGCTTCTCACTCAG 1066
QY 1066 GAGTGAACCTCAGATGTGCTACAGACTCAACTGTCTGGGACATGACAGGGCTTCT 1125
  |||
Db 1067 GAGTGAACCTCAGATGTGCTACAGACTCAACTGTCTGGGACATGACAGGGCTTCT 1126
QY 1126 GAGTGTTCAGCCACTTAAAGCTCGTGTAAAGATCCCTCTAGACTCAGAGAAAT 1185
  |||
Db 1127 GAGTGTTCAGCCACTTAAAGCTCGTGTAAAGATCCCTCTAGACTCAGAGAAAT 1186
QY 1186 GCCACAGAAATGTGAGAGAGCTTGGAGGCTTCTGTGAGAGATGAGGCACTATTGG 1245
  |||
Db 1187 GCCACAGAAATGTGAGAGAGCTTGGAGGCTTCTGTGAGAGATGAGGCACTATTGG 1246
QY 1246 GGAATTTGAGGAGACAGCTTAACTGAGCTGGCTGTGATGTTTGTGACAGTGAACCA 1305
  |||
Db 1247 GGAATTTGAGGAGACAGCTTAACTGAGCTGGCTGTGATGTTTGTGACAGTGAACCA 1306
QY 1306 CAGTGGAGAGAGTTTTCAGTCTGATCTGATCTTAAACAC 1349
  |||
Db 1307 CAGTGGAGAGAGTTTTCAGTCTGATCTGATCTTAAACAC 1350
  |||

RESULT 4
ID AA158658 standard; cDNA, 1288 BP.
XX
XX AA158658;
XX
XX 22-OCT-2001 (first entry)
XX
XX Human polynucleotide SEQ ID NO 861.
XX
XX Human, noctropic; immunosuppressant; cyostatic; gene therapy; cancer;
XX peripheral nervous system; neuropathy; central nervous system; CNS;
XX Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
XX amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
XX chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
XX leukaemia; ss.

```

```

XX OS Homo sapiens.
XX PN WO200153312-A1.
XX 26-JUL-2001.
XX 26-DEC-2000; 2000MO-US34263.
XX 21-JAN-2000; 2000US-0488725.
XX 25-APR-2000; 2000US-0552317.
XX 09-JUL-2000; 2000US-0598042.
XX 19-JUL-2000; 2000US-0620312.
XX 03-AUG-2000; 2000US-0653450.
XX 14-SEP-2000; 2000US-0662191.
XX 19-OCT-2000; 2000US-0693036.
XX 29-NOV-2000; 2000US-0727344.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D,
XX Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J,
XX Zhao QA, Zhou P, Goodrich R, Drmanac RT,
XX
XX WPI: 2001-442253/47.
XX P-PSDB; AAM39502.
XX
XX Novel nucleic acids and polypeptides, useful for treating disorders
XX such as central nervous system injuries -
XX
XX Claim 1; SEQ ID NO 861; 1007bp; English.
XX
XX The invention relates to human nucleic acids (AA157798-AA161369) and
XX the encoded polypeptides (AAM38642-AA42213) with noctropic,
XX immunosuppressant and cyostatic activity. The polynucleotides are useful
XX in gene therapy. A composition containing a polypeptide or polynucleotide
XX of the invention may be used to treat diseases of the peripheral nervous
XX system, such as peripheral nervous injuries, peripheral neuropathy and
XX localised neuropathies and central nervous system diseases, such as
XX Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
XX lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
XX utilisation of the activities such as: Immune system suppression.
XX Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
XX and thrombolytic activity, cancer diagnosis and therapy, drug screening,
XX assays for receptor activity, arthritis and inflammation, leukaemia and
XX C.N.S disorders.
XX Note: The sequence data for this patent did not form part of the printed
XX specification.
XX
XX Sequence 1288 BP; 275 A; 361 C; 418 G; 234 T; 0 other;
XX
XX
XX Query Match 57.4%; Score 932; DB 22; Length 1288;
XX Best Local Similarity 100.0%; Pred. No. 0;
XX Matches 932; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 46 GGGGAAGGGGCGGCGCGGAGACCCGATGCGCGGAGCGGAGCCGAGATGCTTGGCG 105
  |||
Db 47 GGGGAAGGGGCGGCGCGGAGACCCGATGCGCGGAGCGGAGCCGAGATGCTTGGCG 106
QY 106 GGAAGGCGAGACTGCGAGAGCCCGCGGAGCGGAGCCGAGATGCTTGGCG 165
  |||
Db 107 GGAAGGCGAGACTGCGAGAGCCCGCGGAGCGGAGCCGAGATGCTTGGCG 166
QY 166 CTGATAGGGGTGAGCGGCGGACCTGCGAGCGGGAAGTGCACCGTGTGTAAGATCATG 225
  |||
Db 167 CTGATAGGGGTGAGCGGCGGACCTGCGAGCGGGAAGTGCACCGTGTGTAAGATCATG 226
QY 226 GAGTGTCTGGGACAGACGAGTGTGAACGCGGACGCGGGAAGTGTCTTGAAGCCAG 285
  |||
Db 227 GAGTGTCTGGGACAGACGAGTGTGAACGCGGACGCGGGAAGTGTCTTGAAGCCAG 286
  |||
QY 286 GACAGTTCTTAAAGAGTCTGACGCGGAGAGCAAGGCGCTTGAAGAGACGTAC 345
  |||

```

Db 287 GACAGTTCCTACCAAGTCTCTGACGGCAGAGCAGAAAGCCCTTGAAGACAGTAC 346
Qy 346 AATTTGACCATCCAGATGCTTTGATATGATTTGATGACAGAGACTCTGAGAACATC 405
Db 347 AATTTGACCATCCAGATGCTTTGATATGATTTGATGACAGAGACTCTGAGAACATC 406
Qy 406 GTGAGGGCAAAACGGTSGAGGTGCGACCTATGATTTTGTGACACTCAAGGTTACCA 465
Db 407 GTGAGGGCAAAACGGTSGAGGTGCGACCTATGATTTTGTGACACTCAAGGTTACCA 466
Qy 466 GAGACCAAGGTGTGTACCTTCGCGACGTGGTCTGTTGAGGGATCTTGGTGTTCAC 525
Db 467 GAGACCAAGGTGTGTACCTTCGCGACGTGGTCTGTTGAGGGATCTTGGTGTTCAC 526
Qy 526 AGCCAGAGATCCGGGACATGTTCCACCTGCGCTCTTCGTGAGACACCGACCTCCGAGCTC 585
Db 527 AGCCAGAGATCCGGGACATGTTCCACCTGCGCTCTTCGTGAGACACCGACCTCCGAGCTC 586
Qy 586 AGGCTGTCTGAAAGATTCTCCGGGACGTGCGCCGAGGAGGAGACCTTGAGCAGATTCTG 645
Db 587 AGGCTGTCTGAAAGATTCTCCGGGACGTGCGCCGAGGAGGAGACCTTGAGCAGATTCTG 646
Qy 646 AGCGATACACCATCTTGTGTGAAGCCGCTTCGAGGAGTTCTGCTCCGACAAAGAG 705
Db 647 AGCGATACACCATCTTGTGTGAAGCCGCTTCGAGGAGTTCTGCTCCGACAAAGAG 706
Qy 706 TATGCCGATGTGATCATCCACGAGAGTGGACAAATATGTTGCCATCAACTGATCGTG 765
Db 707 TATGCCGATGTGATCATCCACGAGAGTGGACAAATATGTTGCCATCAACTGATCGTG 766
Qy 766 CAGCAGATCCAGAGCATTTCTGAATGTGACATCTTGCAATGGACCGAGAGGGTCCAT 825
Db 767 CAGCAGATCCAGAGCATTTCTGAATGTGACATCTTGCAATGGACCGAGAGGGTCCAT 826
Qy 826 GGGGGGAGCTCAAGCGGACCTTTTCTGAGCCAGGGGACCACTCTGGATGCTGACCTCT 885
Db 827 GGGGGGAGCTCAAGCGGACCTTTTCTGAGCCAGGGGACCACTCTGGATGCTGACCTCT 886
Qy 886 GGCAAAACGGTCAATTTGAGTCCAGCAGACAGACCCCACTAGGGGGTGGCGAGCTCAG 945
Db 887 GGCAAAACGGTCAATTTGAGTCCAGCAGACAGACCCCACTAGGGGGTGGCGAGCTCAG 946
Qy 946 GGCAAGTCTCCCGCCGGCATGTGTTCAGG 977
Db 947 GGCAAGTCTCCCGCCGGCATGTGTTCAGG 978

RESULT 5
AAL36439
ID AAL36439 standard; DNA; 9732 BP.
XX
AC AAL36439;
XX
DT 08-JAN-2002 (first entry)
XX
DE Human musculoskeletal system related polynucleotide SEQ ID NO 2804.
XX
XX Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KM antifungal; hepatotropic; antidiabetic; antiinflammatory; antitumor;
KM vulnerrary; anticonvulsant; antibacterial; antitumor; antiparasitic;
KM cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
KM neurological disease; infection; human; secreted protein;
KM musculoskeletal system; ds.
XX
OS Homo sapiens.
XX
PN WO200155367-A1.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01338.
XX
PR 31-JAN-2000; 2000US-0179065.

PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225759.
PR 14-AUG-2000; 2000US-0225759.
PR 14-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226688.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 06-SEP-2000; 2000US-0231242.
PR 06-SEP-2000; 2000US-0231243.
PR 06-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.

PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 17-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249284.
 PR 17-NOV-2000; 2000US-0249285.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251889.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Barash SC, Ruben SM;
 XX
 DR WPI; 2001-451937/48.
 XX
 PT Isolated polypeptide for treating, preventing and/ or prognosing
 PT disorders related to the musculoskeletal system including
 PT musculoskeletal cancers and also for testing and detection e.g.
 PT diagnosis -

XX
 PS Example 2; SEQ ID NO 2804; 781bp + Sequence Listing; English.
 XX
 CC The invention relates to novel genes (AAL34669-AAL37666) and proteins
 CC (AAB03087-ABB04109) associated with the musculoskeletal system useful
 CC for preventing, treating or ameliorating medical conditions e.g. by
 CC protein or gene therapy. The genes are isolated from a range of human
 CC tissues disclosed in the specification. The nucleic acids, proteins,
 CC antibodies and (ant)agonists are useful in the diagnosis, treatment,
 CC and prevention of: (a) cancer, e.g. breast and ovarian cancer and
 CC other cancers of the adrenal gland, bone, bone marrow, breast,
 CC gastrointestinal tract, liver, lung, or urogenital; (b) immune
 CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
 CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
 CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis;
 CC (c) cardiovascular disorders such as myocardial ischaemia; (d) wound
 CC healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy;
 CC and (f) infectious diseases such as viral, bacterial, fungal and
 CC parasitic infections.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIP0 at ftp.wipo.int/pub/published_pct_sequences.
 XX

SQ Sequence 9732 BP; 2063 A; 2805 C; 2561 G; 2303 T; 0 other;

Query Match 54.2%; Score 880; DB 22; Length 9732;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 880; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 745 GTTGCCATCAACCTGATGTCAGCATCCAGGACATTCGATGATGATGATCTGCAAA 804
 DB 3901 GTTGCCATCAACCTGATGTCAGCATCCAGGACATTCGATGATGATGATCTGCAAA 3960
 QY 805 TGGCACCAGAGAGAGGTTCATGAGCGAGCTCAACAGGACCTTTCTGAGCCAGGGAGC 864
 DB 3961 TGGCACCAGAGAGAGGTTCATGAGCGAGCTCAACAGGACCTTTCTGAGCCAGGGAGC 4020
 QY 865 CACCTTGGATGATGATCTCTGCAAAACGCTCATATTGATGATGATGATGATGATGATGAT 924
 DB 4021 CACCTTGGATGATGATCTCTGCAAAACGCTCATATTGATGATGATGATGATGATGATGAT 4080
 QY 925 TGAGGGGCTGCGAGGCTCAGGGCAGGCTCCCGCCGCGATGATGATGATGATGATGATGAT 984
 DB 4081 TGAGGGGCTGCGAGGCTCAGGGCAGGCTCCCGCCGCGATGATGATGATGATGATGATGAT 4140
 QY 985 CTTGGGAGCGCCACCCACCACTGCTCTCTGCGGCAACCCAGGGAGATGTTAGC 1044
 DB 4141 CTTGGGAGCGCCACCCACCACTGCTCTCTGCGGCAACCCAGGGAGATGTTAGC 4200
 QY 1045 AGCGAGGCTTCTCTCACTAGAGATGGAATCAATGATGATGATGATGATGATGATGATGAT 1104
 DB 4201 AGCGAGGCTTCTCTCACTAGAGATGGAATCAATGATGATGATGATGATGATGATGATGAT 4260
 QY 1105 GGGACACTGACAGGGCTCTGAGGTTTCAGCACTTAGGCTCGTTGGGTTTAAAGAT 1164
 DB 4261 GGGACACTGACAGGGCTCTGAGGTTTCAGCACTTAGGCTCGTTGGGTTTAAAGAT 4320
 QY 1165 CCCTTAGGTCATCTGGAATGCAAGATGTCAGAGAGCTGGAGGCTTCTGTGAG 1224
 DB 4321 CCCTTAGGTCATCTGGAATGCAAGATGTCAGAGAGCTGGAGGCTTCTGTGAG 4380
 QY 1225 GAATGTGAGGACATTTATGGGAAATTTAGAGAGACGCTTGACACTGCTGGCTGAT 1284
 DB 4381 GAATGTGAGGACATTTATGGGAAATTTAGAGAGAGCTTGACACTGCTGGCTGAT 4440
 QY 1285 GTTTTGTGACAGTGAACCAAGTGGAGAGAGTTTTCAGTGTGATGATGATGATGATGATGAT 1344
 DB 4441 GTTTTGTGACAGTGAACCAAGTGGAGAGAGTTTTCAGTGTGATGATGATGATGATGATGAT 4500
 QY 1345 CACACTCACACATTAACCAAAAGTTTGTGAACAAGTACTTCTTTTACATGTTA 1404
 DB 4501 CACACTCACACATTAACCAAAAGTTTGTGAACAAGTACTTCTTTTACATGTTA 4560

Qy 1405 CATGTCCTCATGTTTCTGTTTCTGTTTCATTAACACAAAGCTGGTTGGCCTTAACAAC 1464
|||
Db 4561 CATGTCCTCATGTTTCTGTTTCTGTTTCTAATACAAAGCTGGTTGGCCTTAACAAC 4620
|||
Qy 1465 CTAATTTATGACCCAGTGTGTTTGACATTCACAGCGTGCCTTACAGGATATGGGAGCCAC 1524
|||
Db 4621 CTAATTTATGACCCAGTGTGTTTGACATTCACAGCGTGCCTTACAGGATATGGGAGCCAC 4680
|||
Qy 1525 TGAGGAGATGTTTCCCGCTTGCTGTCCTTAAGGAGGAGGAGGAGGAGGAGTGCCT 1584
|||
Db 4681 TGAGGAGATGTTTCCCGCTTGCTGTCCTTAAGGAGGAGGAGGAGGAGGAGTGCCT 4740
|||
Qy 1585 GGAAGCACCCAGCATCACACCCAGGCTTGTGCGGGCCAG 1624
|||
Db 4741 GGAAGCACCCAGCATCACACCCAGGCTTGTGCGGGCCAG 4780
|||
RESULT 6
ABX59427
ID ABX59427 standard; cDNA; 9732 BP.
XX
XX ABX59427;
AC
XX 26-FEB-2003 (first entry)
DT
XX
DE cDNA encoding novel human musculoskeletal system antigen #1771.
XX
XX Gene; 68; musculoskeletal system antigen; cancer; metastasis;
KM re-vascularisation; thrombosis; arteriosclerosis; mineral content;
KM cardiovascular condition; wound; injury; burn; angiogenesis; ulcer;
KM post-operative tissue repair; limb regeneration; neuronal growth;
KM neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
KM AIDS-related complex; chondrocyte growth; bone regeneration;
KM periodontal regeneration; tissue transport; bone graft; skin aging;
KM keratinocyte growth; hair loss; melanocyte growth; cell proliferation;
KM cell growth; organ transplant; cell differentiation; body height;
KM weight; hair colour; eye colour; skin; percentage of adipose tissue;
KM pigmentation; cosmetic surgery; metabolism; biorythm; cardiac rhythm;
KM depression; tendency for violence; pain; reproductive capability;
KM hormone level; endocrine level; appetite; libido; memory; stress;
KM storage capability; fat content; lipid content; protein content;
KM carbohydrate content; vitamin content; cofactor content;
KM nutritional component.
XX
XX Homo sapiens.
OS
XX
XX US2002147140-A1.
PN
XX 10-OCT-2002.
PD
XX
XX 17-JAN-2001; 2001US-0764877.
PF
XX
XX 31-JAN-2000; 2000US-179065P.
PR 04-FEB-2000; 2000US-180628P.
PR 28-JUN-2000; 2000US-214886P.
PR 07-JUL-2000; 2000US-216647P.
PR 27-JUL-2000; 2000US-216880P.
PR 11-JUL-2000; 2000US-217487P.
PR 11-JUL-2000; 2000US-217496P.
PR 14-JUL-2000; 2000US-218290P.
PR 26-JUL-2000; 2000US-220963P.
PR 26-JUL-2000; 2000US-220964P.
PR 14-AUG-2000; 2000US-224518P.
PR 14-AUG-2000; 2000US-224519P.
PR 14-AUG-2000; 2000US-225267P.
PR 14-AUG-2000; 2000US-225268P.
PR 14-AUG-2000; 2000US-225470P.
PR 14-AUG-2000; 2000US-225471P.
PR 14-AUG-2000; 2000US-225575P.
PR 14-AUG-2000; 2000US-225758P.
PR 22-AUG-2000; 2000US-226868P.
PR 30-AUG-2000; 2000US-228924P.
PR 01-SEP-2000; 2000US-229287P.

PR 01-SEP-2000; 2000US-229343P.
PR 01-SEP-2000; 2000US-229344P.
PR 01-SEP-2000; 2000US-229345P.
PR 05-SEP-2000; 2000US-229509P.
PR 05-SEP-2000; 2000US-229513P.
PR 08-SEP-2000; 2000US-231413P.
PR 21-SEP-2000; 2000US-234223P.
PR 21-SEP-2000; 2000US-234274P.
PR 25-SEP-2000; 2000US-234997P.
PR 27-SEP-2000; 2000US-235834P.
PR 29-SEP-2000; 2000US-236327P.
PR 29-SEP-2000; 2000US-236367P.
PR 29-SEP-2000; 2000US-236368P.
PR 29-SEP-2000; 2000US-236369P.
PR 29-SEP-2000; 2000US-236370P.
PR 02-OCT-2000; 2000US-236802P.
PR 02-OCT-2000; 2000US-237037P.
PR 02-OCT-2000; 2000US-237038P.
PR 02-OCT-2000; 2000US-237039P.
PR 02-OCT-2000; 2000US-237040P.
PR 13-OCT-2000; 2000US-239935P.
PR 20-OCT-2000; 2000US-240960P.
PR 20-OCT-2000; 2000US-241785P.
PR 20-OCT-2000; 2000US-241809P.
PR 01-NOV-2000; 2000US-244617P.
PR 17-NOV-2000; 2000US-249299P.
PR 08-DEC-2000; 2000US-251856P.
PR 08-DEC-2000; 2000US-251868P.
PR 08-DEC-2000; 2000US-251869P.
XX
XX (ROSEN/ ROSEN C A.
PA (RUBE/ RUBEN S M.
PA (BARA/ BARASH S C.
XX
XX Rosen CA, Ruben SM, Barash SC;
PI
XX
XX WPI; 2003-128199/12.
DR
XX
XX Isolated nucleic acid molecules encoding musculoskeletal system
PT associated polypeptides, useful for detecting disorders, e.g. cancer -
PT
XX
XX Disclosure; SEQ ID NO 2804; 321bp; English.
XX
XX The invention describes an isolated nucleic acid molecule comprising a
CC sequence encoding musculoskeletal system associated polypeptides useful
CC for detecting disorders, e.g., cancer or cancer metastases, in animals
CC or humans. The nucleic acid; stimulates re-vascularisation of ischaemic
CC tissues associated with conditions such as thrombosis, arteriosclerosis,
CC and other cardiovascular conditions; treats wounds due to injuries,
CC burns, post-operative tissue repair, and ulcers; stimulates angiogenesis
CC and limb regeneration; stimulates neuronal growth; can treat and prevent
CC neuronal damage occurring in certain disorders or neurodegenerative
CC conditions, such as, Alzheimer's disease, Parkinson's disease, and
CC AIDS-related complex; stimulates chondrocyte growth, thus they can be
CC used to enhance bone and periodontal regeneration and aid in tissue
CC transports or bone grafts; prevents skin aging due to sunburn by
CC stimulating keratinocyte growth; prevents hair loss, since FGF family
CC members activate hair-forming cells and promotes melanocyte growth;
CC stimulates growth and differentiation of hematopoietic cells and bone
CC marrow cells when used in combination with other cytokines; maintains
CC organs before transplantation or for supporting cell culture of primary
CC tissues; induces tissue of mesodermal origin to differentiate in early
CC embryos; increases or decreases the differentiation or proliferation of
CC embryonic stem cells, besides, hematopoietic lineage; modulates
CC mammalian characteristics, such as, body height, weight, hair colour, eye
CC colour, skin, percentage of adipose tissue, pigmentation, size, and shape
CC (e.g., cosmetic surgery); modulates mammalian metabolism; changes
CC mammal's metal state or physical state by influencing biorythms;
CC cardiac rhythms, depression, tendency for violence, tolerance for pain,
CC reproductive capabilities, hormonal or endocrine levels, appetite,
CC libido, memory, or stress; increases or decreases storage capabilities,
CC fat content, lipid, protein, carbohydrate, vitamins, minerals, cofactors
CC or other nutritional components. This sequence encodes a novel human

CC musculoskeletal system antigen.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from the US patent office at
CC ftp.segdata.uspto.gov/sequence.html?DocID=20020147140.
CC
XX
SQ Sequence 9732 BP; 2063 A; 2805 C; 2561 G; 2303 T; 0 other;

Query Match 54.2%; Score 880; DB 25; Length 9732;

Best Local Similarity 100.0%; Pred. No. 0;
Matches 880; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 745 GTTCCATCAACCTGATGTCAGACACATCCAGGACATTTGAAATGAGCATCTGCAAA 804
Db 3901 GTTCCATCAACCTGATGTCAGACACATCCAGGACATTTGAAATGAGCATCTGCAAA 3960
QY 805 TGGCACCAGAGAGGGTCCATGAGCGGAGCTACAAAGCGACCTTTCTGAGCCAGGGAG 864
Db 3961 TGGCACCAGAGAGGGTCCATGAGCGGAGCTACAAAGCGACCTTTCTGAGCCAGGGAG 4020
QY 865 CACCTTGAGATGTCGACCTCTGAGCAACGGTCACATTTGAGTCCAGACGACCCAC 924
Db 4021 CACCTTGAGATGTCGACCTCTGAGCAACGGTCACATTTGAGTCCAGACGACCCAC 4080
QY 925 TGAGGGGCTGCGGAGCTCAGGGAGGTCCTCCGCGCGGACATGTTGTTGAGGGACTGAG 984
Db 4081 TGAGGGGCTGCGGAGCTCAGGGAGGTCCTCCGCGCGGACATGTTGTTGAGGGACTGAG 4140
QY 985 CTTGGGAGCGCCCAACCCACACCCACTGTTCTCTCGGCGGACCCAGGGAGGTTAGC 1044
Db 4141 CTTGGGAGCGCCCAACCCACACCCACTGTTCTCTCGGCGGACCCAGGGAGGTTAGC 4200
QY 1045 AGCGAGGCTTCTCTCAGTCACTGAGGATGGAATCACTGATGTCATCTCAACTGCT 1104
Db 4201 AGCGAGGCTTCTCTCAGTCACTGAGGATGGAATCACTGATGTCATCTCAACTGCT 4260
QY 1105 GGGACATGACAGGCGTTCCTGAGGTTTTCAGCCACTTAAAGCTGTTGCGGTTAAAGAT 1164
Db 4261 GGGACATGACAGGCGTTCCTGAGGTTTTCAGCCACTTAAAGCTGTTGCGGTTAAAGAT 4320
QY 1165 CCCTCTAGGTCAGTGAAGATGCCACAGATGTCAGGAGCCCTGGAGGCTTCTGTAG 1224
Db 4321 CCCTCTAGGTCAGTGAAGATGCCACAGATGTCAGGAGCCCTGGAGGCTTCTGTAG 4380
QY 1225 GAATGTAGGCACTTATTTGGGGAAATTGAGGACACGCTTACACATGCTGCTGAT 1284
Db 4381 GAATGTAGGCACTTATTTGGGGAAATTGAGGACACGCTTACACATGCTGCTGAT 4440
QY 1285 GTTTGTTGACAGTGAACCCACAGTGGGAGAGATTTTTCAGTGCATGCTGCTTCTTA 1344
Db 4441 GTTTGTTGACAGTGAACCCACAGTGGGAGAGATTTTTCAGTGCATGCTGCTTCTTA 4500
QY 1345 CACACTCACACATTAATCTCAAAAGTTTGTGAACAAATCTTCTTTTAACTGTTA 1404
Db 4501 CACACTCACACATTAATCTCAAAAGTTTGTGAACAAATCTTCTTTTAACTGTTA 4560
QY 1405 CATGCTCTCATGTTTCTGTTTCTGTTTCAATACACAGGCTGTGTGCTCTCAAAAC 1464
Db 4561 CATGCTCTCATGTTTCTGTTTCTGTTTCAATACACAGGCTGTGTGCTCTCAAAAC 4620
QY 1465 CTAAATTCATGACCCAGTGTGTTGCTGAGTCCAGGCTTACAGATATGAGGAGCCAC 1524
Db 4621 CTAAATTCATGACCCAGTGTGTTGCTGAGTCCAGGCTTACAGATATGAGGAGCCAC 4680
QY 1525 TGAAGGATGTTTCCCTTGTGCTTAAAGCAGAGGAGCCAGATGCTT 1584
Db 4681 TGAAGGATGTTTCCCTTGTGCTTAAAGCAGAGGAGCCAGATGCTT 4740
QY 1585 GGAAGCACCAGCATCACCCAGGCTTGTGCGGGGCGAG 1624
Db 4741 GGAAGCACCAGCATCACCCAGGCTTGTGCGGGGCGAG 4780
```

```
RESULT 7
ID AAL36440
AC AAL36440;
XX AAL36440;
DT 08-JAN-2002 (first entry)
XX
DE Human musculoskeletal system related polynucleotide SEQ ID NO 2805.
XX
KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor;
KW vulnereary; anticoagulant; antibacterial; antifungal; antiparasitic;
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW neurological disease; infection; human; secreted protein;
KW musculoskeletal system; ds.
XX
OS Homo sapiens.
XX
PN NO200155367-A1.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001MO-US01338.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226688.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
```

PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0233081.
PR 12-SEP-2000; 2000US-0233968.
PR 14-SEP-2000; 2000US-0233397.
PR 14-SEP-2000; 2000US-0233398.
PR 14-SEP-2000; 2000US-0233399.
PR 14-SEP-2000; 2000US-0234400.
PR 14-SEP-2000; 2000US-0234401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.

PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-451937/48.
XX
XX Isolated polypeptide for treating, preventing and/or prognosing
PT disorders related to the musculoskeletal system including
PT musculoskeletal cancers and also for testing and detection e.g.
PT diagnosis -
XX
PS Example 2; SEQ ID NO 2805; 781bp + Sequence Listing; English.
XX
XX The invention relates to novel genes (AAL34669-AAL37666) and proteins
CC (AAB01087-AAB04109) associated with the musculoskeletal system useful
CC for preventing, treating or ameliorating medical conditions e.g. by
CC protein or gene therapy. The genes are isolated from a range of human
CC tissues disclosed in the specification. The nucleic acids, proteins,
CC antibodies and (ant)agonists are useful in the diagnosis, treatment,
CC and prevention of: (a) cancer, e.g. breast and ovarian cancer and
CC other cancers of the adrenal gland, bone, bone marrow, breast,
CC gastrointestinal tract, liver, lung, or urogenital; (b) immune
CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis;
CC (c) cardiovascular disorders such as myocardial ischaemia; (d) wound
CC healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy;
CC and (f) infectious diseases such as viral, bacterial, fungal and
CC parasitic infections.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pat_sequences.
XX
SQ Sequence 19125 BP; 4263 A; 5265 C; 5444 G; 4153 T; 0 other;
Query Match 54.2%; Score 880; DB 22; Length 19125;
Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Gaps 0;
Matches 880; Conservative 0; Indels 0; Gaps 0;
QY 745 GTTGCCATCAACCTGATGTCGAGCAATCCAGAGATTCTGAATGTCATCTGCATAA 804
DB 5997 GTTGCCATCAACCTGATGTCGAGCAATCCAGAGATTCTGAATGTCATCTGCATAA 6056
QY 805 TGGCACCAGAGAGGTCCTGAGCAATGAGCGAGCTTAAAGCGAATCTTTCTGAGCGAGGAC 864
DB 6057 TGGCACCAGAGAGGTCCTGAGCAATGAGCGAGCTTAAAGCGAATCTTTCTGAGCGAGGAC 6116
QY 865 CACCCTGGAGTGTGCACTCTGCAAAAGGTCACTTTGGAGTCGAGCAGACCCAC 924
DB 6117 CACCCTGGAGTGTGCACTCTGCAAAAGGTCACTTTGGAGTCGAGCAGACCCAC 6176
QY 925 TGAAGGGCTGCCAGCTTCAGGCGAGGCTCCCGCCGAGCATGTGTTCAGGAGCTGAG 984
DB 6177 TGAAGGGCTGCCAGCTTCAGGCGAGGCTCCCGCCGAGCATGTGTTCAGGAGCTGAG 6236
QY 985 CTTGGGGAGCGCCACCAACCAACCACTGCTTCTCTGCGCGCAGCCCAAGGGAGTGTAGC 1044


```

Dh 6237 CTTGGGAGCCGCCACCCACACCCACTGCTTCTCTCGGCGCACCCGAGGAGTGTAGC 6296
Qy 1045 AGGAGGSCCTTCTCACTAGAGTGAATCTGAGTGTCACTGACTCAACTTGGCT 1104
Dh 6297 AGCGAGGCTTCTCACTAGAGTGAATCTGAGTGTCACTGACTCAACTTGGCT 6356
Qy 1105 GGGACACTGACAGGCGCTTCTGAGGTTTTCAGCCACTTAGGCTGCTGGCTTTAAAGAT 1164
Dh 6357 GGGACACTGACAGGCGCTTCTGAGGTTTTCAGCCACTTAGGCTGCTGGCTTTAAAGAT 6416
Qy 1165 CCCTCTAGTCACTGAGAAATGCCACAAATGTGCAGAAAGCTGGAGGCTTCTGTAG 1224
Dh 6417 CCCTCTAGTCACTGAGAAATGCCACAAATGTGCAGAAAGCTGGAGGCTTCTGTAG 6476
Qy 1225 GAATGTAGGACACTTATTGGGGAAATTTGAGAGACAGCCTTAGCACTGGCTGGCCGAT 1284
Dh 6477 GAATGTAGGACACTTATTGGGGAAATTTGAGAGACAGCCTTAGCACTGGCTGGCCGAT 6536
Qy 1285 GTTTTGTGACAGTGAACCAACAGTGGAGAGATTTTTCAGTCTGATCTGGTCTTA 1344
Dh 6537 GTTTTGTGACAGTGAACCAACAGTGGAGAGATTTTTCAGTCTGATCTGGTCTTA 6596
Qy 1345 CACACTGACACACTAACTCAAAAGTTTGTGAACAAATCTTCTTTTACATGTTA 1404
Dh 6597 CACACTGACACACTAACTCAAAAGTTTGTGAACAAATCTTCTTTTACATGTTA 6656
Qy 1405 CAGTCTCACTGTTTCTGTTTCTGTTTCTATTAACAAGGCTGTGTGGCTTAACAAC 1464
Dh 6657 CAGTCTCACTGTTTCTGTTTCTGTTTCTATTAACAAGGCTGTGTGGCTTAACAAC 6716
Qy 1465 CTATTTTCAAGACCCAGTGTGTTGTCAGTCCAGCGTGGCTTAACAAGTATGGGAGGCAC 1524
Dh 6717 CTATTTTCAAGACCCAGTGTGTTGTCAGTCCAGCGTGGCTTAACAAGTATGGGAGGCAC 6776
Qy 1525 TGAAGGAGTGTTCCTCCCTGCTTGTGCTTAAAGCAGAGAGCGAGATGCCCT 1584
Dh 6777 TGAAGGAGTGTTCCTCCCTGCTTGTGCTTAAAGCAGAGAGCGAGATGCCCT 6836
Qy 1585 GGAAGACCCAGATCAACCCAGGCTTGTGCGGGGCGAG 1624
Dh 6837 GGAAGACCCAGATCAACCCAGGCTTGTGCGGGGCGAG 6876

```

RESULT 8

ABX59428

ID ABX59428 standard; cDNA; 19125 BP.

XX AC ABX59428;

DT 26-FEB-2003 (first entry)

DE cDNA encoding novel human musculoskeletal system antigen #1772.

XX Gene; ss; musculoskeletal system antigen; cancer; metastasis;
 XX re-vascularisation; thrombosis; arteriosclerosis; mineral content;
 XX cardiovascular condition; wound; injury; burn; angiogenesis; ulcer;
 XX post-operative tissue repair; limb regeneration; neuronal growth;
 XX neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
 XX AIDS-related complex; chondrocyte growth; bone regeneration;
 XX periodontal regeneration; tissue transport; bone graft; skin aging;
 XX keratinocyte growth; hair loss; melanocyte growth; cell proliferation;
 XX cell growth; organ transplant; cell differentiation; body height;
 XX weight; hair colour; eye colour; skin; percentage of adipose tissue;
 XX pigmentation; cosmetic surgery; metabolism; biorythm; cardiac rhythm;
 XX depression; tendency for violence; pain; reproductive capability;
 XX hormone level; endocrine level; appetite; libido; memory; stress;
 XX storage capability; fat content; lipid content; protein content;
 XX carbohydrate content; vitamin content; cofactor content;
 XX nutritional component.

XX Homo sapiens.

XX OS US2002147140-A1.

```

XX 10-OCT-2002.
PD 17-JAN-2001; 2001US-0764877.
XX 31-JAN-2000; 2000US-179065P.
PR 04-FEB-2000; 2000US-180628P.
PR 28-JUN-2000; 2000US-214886P.
PR 07-JUL-2000; 2000US-216647P.
PR 07-JUL-2000; 2000US-216880P.
PR 11-JUL-2000; 2000US-217487P.
PR 11-JUL-2000; 2000US-217496P.
PR 14-JUL-2000; 2000US-218290P.
PR 26-JUL-2000; 2000US-220963P.
PR 26-JUL-2000; 2000US-220964P.
PR 14-AUG-2000; 2000US-224518P.
PR 14-AUG-2000; 2000US-224519P.
PR 14-AUG-2000; 2000US-225267P.
PR 14-AUG-2000; 2000US-225268P.
PR 14-AUG-2000; 2000US-225270P.
PR 14-AUG-2000; 2000US-225447P.
PR 14-AUG-2000; 2000US-225757P.
PR 14-AUG-2000; 2000US-225758P.
PR 22-AUG-2000; 2000US-226868P.
PR 30-AUG-2000; 2000US-228924P.
PR 01-SEP-2000; 2000US-229287P.
PR 01-SEP-2000; 2000US-229343P.
PR 01-SEP-2000; 2000US-229344P.
PR 01-SEP-2000; 2000US-229345P.
PR 05-SEP-2000; 2000US-229509P.
PR 05-SEP-2000; 2000US-229513P.
PR 08-SEP-2000; 2000US-231413P.
PR 21-SEP-2000; 2000US-234223P.
PR 21-SEP-2000; 2000US-234274P.
PR 25-SEP-2000; 2000US-234997P.
PR 27-SEP-2000; 2000US-235834P.
PR 29-SEP-2000; 2000US-236337P.
PR 29-SEP-2000; 2000US-236357P.
PR 29-SEP-2000; 2000US-236368P.
PR 29-SEP-2000; 2000US-236369P.
PR 29-SEP-2000; 2000US-236370P.
PR 02-OCT-2000; 2000US-236802P.
PR 02-OCT-2000; 2000US-237037P.
PR 02-OCT-2000; 2000US-237038P.
PR 02-OCT-2000; 2000US-237039P.
PR 02-OCT-2000; 2000US-237040P.
PR 13-OCT-2000; 2000US-239935P.
PR 20-OCT-2000; 2000US-240960P.
PR 20-OCT-2000; 2000US-241785P.
PR 20-OCT-2000; 2000US-241809P.
PR 01-NOV-2000; 2000US-244617P.
PR 17-NOV-2000; 2000US-249299P.
PR 08-DEC-2000; 2000US-251856P.
PR 08-DEC-2000; 2000US-251868P.
PR 08-DEC-2000; 2000US-251869P.
XX (ROSE/) ROSEN C A.
PA (RUBE/) RUBEN S M.
PA (BARA/) BARASH S C.
XX Rosen CA, Ruben SM, Barash SC,
DR WPI, 2003-128199/12.
XX Isolated nucleic acid molecules encoding musculoskeletal system
PT associated polypeptides, useful for detecting disorders, e.g. cancer -
XX Disclosure; SEQ ID NO 2805; 321pp; English.
XX The invention describes an isolated nucleic acid molecule comprising a
CC sequence encoding musculoskeletal system associated polypeptides useful
CC for detecting disorders, e.g., cancer or cancer metastases, in animals
CC or humans. The nucleic acid: stimulates re-vascularisation of ischaemic

```

tissues associated with conditions such as thrombosis, arteriosclerosis, and other cardiovascular conditions; treats wounds due to injuries, burns, post-operative tissue repair, and ulcers; stimulates angiogenesis and limb regeneration; stimulates neuronal growth; can treat and prevent neuronal damage occurring in certain disorders or neurodegenerative conditions, such as, Alzheimer's disease, Parkinson's disease, and AIDS-related complex; stimulates chondrocyte growth, thus they can be used to enhance bone and periodontal regeneration and aid in tissue transports or bone grafts; prevents skin aging due to sunburn by stimulating keratinocyte growth; prevents hair loss, since FGF family members activate hair-forming cells and promotes melanocyte growth; stimulates growth and differentiation of hematopoietic cells and bone marrow cells when used in combination with other cytokines; maintains organs before transplantation or for supporting cell culture of primary tissues; induces tissue of mesodermal origin to differentiate in early embryonic; increases or decreases the differentiation or proliferation of embryonic stem cells; besides, haematopoietic lineage; modulates mammalian characteristics, such as, body height, weight, hair colour, eye colour, skin, percentage of adipose tissue, pigmentation, size, and shape (e.g., cosmetic surgery); modulates mammalian metabolism; changes mammalian's metal state or physical state by influencing biorhythms, circadian rhythms, depression, tendency for violence, tolerance for pain, reproductive capabilities, hormonal or endocrine levels, appetite, libido, memory, or stress; increases or decreases storage capabilities, fat content, lipid, protein, carbohydrate, vitamins, minerals, cofactors or other nutritional components. This sequence encodes a novel human musculoskeletal system antigen.

Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from the US patent office at <http://seqdata.uspto.gov/sequence.html?DocID=20020147140>.

SQ Sequence 19125 BP; 4263 A; 5265 C; 5444 G; 4153 T; 0 other;

Query Match	54.2%;	Score 880;	DB 25;	Length 19125;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 880; Conservative	0;	Mismatches	0;	Gaps 0;

Qy 745 GTTGCATCAACCTGATCGTGACAGACATCCAGGACATTCTGAA TGTGCATCTGCAAA 804

Dh 5987 GTTGCATCAACCTGATCGTGACAGACATCCAGGACATTCTGAA TGTGCATCTGCAAA 8056

805 TGCCACCGAGGAGGCTCAATGGGGGAGCTACAAGCGGCTTTCTGTGAGCCAGGGGAC 864

QY 865 CACCCTGGATGCTGACCTCTGGSCAAACGTCACATTGGAGTCCAGCAGCAGACCCCA 924

925 TGAGGGGCTGCCGAGCTTCAGGGCAGGTCTCCCCGCCGGCATGTGTTCAGGSACTGAG 984

6177	19AGGGGAC18CCGGACCC1AAGGGACAG1TCCCGCCCGGATG1G1G1T1AAGGGAC1GAG	6236
985	CCTGGGGAGCCCAACACACCCCACTGCTTCCTCTCGGGCAACCCAGGGGAGTGTAC	1044

DB 623 CCGGGGACGCCCAACACACCCCACTGCTCTCGGGGCACCCCAAGGAGATGTAGC 6296

QY 1045 AGCGAGGCGCTTCACACTAGAGAGCGAACTCAGATGTCCTACTCAAGTCAACTTGGT 11044

Db 6287 AGCAGGCTTCTCTACCTCAGAGTGGAACTCAGATGTGTCACTCAGACTCAACTTGT 6356

Dy 1105 GGGACACTGACAGGCGCTTCTGAGTTTTCAGCCACTTAAGGCTCGTTCGGGTTTAAAGAT 11644

Dd	6357	GGGACACTGCACGGGCGTTCTCTGAGGTTTTCAGCCACTTTCAGGCTGTTGGGTTTAAAGAT	6416
Dy	1165	CCCTTAGGTCTCATGAGAAATGCCACAGATGTCGAGGAAGCTTCGGAGGCTTCTGTGAG	1224

1225 GAATGTGAGGCACATTA TTGGGGGAATTATGGAACAGGCTTAAACA TTTGGCTGGCTTCAAT 1284
6417 CCTCTAGGTCAC TGAGAAATGCCACAGATGTGAGGAAGCCCTGGAGGCTTCTGTAG 6476

Db 6477 GAATGTGAGGCACATTATTGGGGAATTGAGAGCAGCCTAGACACTGGCTGGCCTGAT 6536

QY	1285	GTTCCTGTTGAAGTGAAGAACCCACAGTGGGAGAGACTTTTCCATCTGATCTGGTCTTA	1344
Db	6537	GTTCCTGTTGAAGTGAAGAACCCACAGTGGGAGAGACTTTTCCATCTGATCTGGTCTTA	6596
QY	1345	CACACTCACACATTAACCTCAAAGTTTGTGAACAAGTACTCTTCCTTTTACATGTTA	1404
Db	6537	CACACTCACACATTAACCTCAAAGTTTGTGAACAAGTACTCTTCCTTTTACATGTTA	6596
QY	1405	CATGTCCTCATGTTTCTGTTTCTGTTTCTTAACAACAAGGCTGTTGTGGCTTACAAC	1464
Db	6657	CATGTCCTCATGTTTCTGTTTCTGTTTCTTAACAACAAGGCTGTTGTGGCTTACAAC	6716
QY	1465	CTAATTTTCATGACCCAGTGTTGCTTCAGTCGACGCTGCTTACACGAGTATGCGGAGCCAC	1524
Db	6717	CTAATTTTCATGACCCAGTGTTGCTTCAGTCGACGCTGCTTACACGAGTATGCGGAGCCAC	6776
QY	1525	TGAGGGATGTTTTCCCTTCTGTTGTGCTTTAAAGCAGAGAACGAGCGGATGCCCC	1584
Db	6777	TGAGGGATGTTTTCCCTTCTGTTGTGCTTTAAAGCAGAGAACGAGCGGATGCCCC	6836
QY	1585	GGAAGCACCCGACATCACACCCAGAGCTTGTGCGGGGCGAC	1624
Db	6837	GGAAGCACCCGACATCACACCCAGAGCTTGTGCGGGGCGAC	6876

RESULT 9	
AAK98735	
ID	AAK98735 standard; DNA; 834 BP.
XX	
AC	AAK98735;

DT 02-MAY-2002 (first entry)

DNA of a human uridine kinase

Human; uridine kinase; diagnostic assay

antibody; vaccine; human ovarian cancer; immunological disorder;

[illegible]

XXC
X

FT	1.780
CDS	1.780

$\partial/\partial T$

XX:

XX	2
	3
	4
	5
	6
	7
	8
	9
	10
	11
	12

XX	2
	6
	8
	9
	10
	11
	12
	13
	14
	15
	16
	17
	18
	19
	20
	21
	22
	23
	24
	25
	26
	27
	28
	29
	30
	31
	32
	33
	34
	35
	36
	37
	38
	39
	40
	41
	42
	43
	44
	45
	46
	47
	48
	49
	50
	51
	52
	53
	54
	55
	56
	57
	58
	59
	60
	61
	62
	63
	64
	65
	66
	67
	68
	69
	70
	71
	72
	73
	74
	75
	76
	77
	78
	79
	80
	81
	82
	83
	84
	85
	86
	87
	88
	89
	90
	91
	92
	93
	94
	95
	96
	97
	98
	99
	100

XX

XX

XX

XX

DR P-PSDB; AA014412.

PT Novel human uridi

PT treating conditions associated with uridine kinase imbalance -

PS Claim 5; Page 22-23; 31pp; English.

CC The invention relates to newly identified human uridine kinase (UDK)

CC polypeptides by recombinant techniques. Also disclosed in the invention


```
QY 54 GGGGGGCGGGGGAACCGATGCGCGGAGCGAGAGCCGAGATGCTTCGGCGGGAGGGA 113
Db 1 GGGGGGCGGGGGAACCGATGCGCGGAGCGAGAGCCGAGATGCTTCGGCGGGAGGGA 60
QY 114 AGACTGCGAGAGCCCGCGCGGAGCCGACCGTCCGACCAAGCGCCCTTCTGATAGG 173
Db 61 AGACTGCGAGAGCCCGCGCGGAGCCGACCGTCCGACCAAGCGCCCTTCTGATAGG 120
QY 174 GGTGAGCGGGGCACTGCGACGGGGAAGTCAGCCGTGTGTAGAAGATCAGAGATGCT 233
Db 121 GGTGAGCGGGGCACTGCGAGGGGAAGTCAGCCGTGTGTAGAAGATCAGAGATGCT 180
QY 234 GGGGACAGACGAGGTGGAACAGCGGCAAGGAGGTGCTCACTGAGCCAGACAGATT 293
Db 181 GGGACAGACGAGGTGGAACAGCGGCAAGGAGGTGCTCACTGAGCCAGACAGATT 240
QY 294 CTACAGAGTCTGACGGGAGAGCAGAGAGCCAGGCTTGAAGAGACATATTTTGA 353
Db 241 CTACAGAGTCTGACGGGAGAGCAGAGAGCCAGGCTTGAAGAGACATATTTTGA 300
QY 354 CCATCCAGATGCTTGTATATGATTTGATGACAGACCTGTAAGAACATCGTGGAGG 413
Db 301 CCATCCAGATGCTTGTATATGATTTGATGACAGACCTGTAAGAACATCGTGGAGG 360
QY 414 CAAACGCTGAGGAGTCCGACCTATGATTTGTGACACACTCAAGTTACAGAGACAC 473
Db 361 CAAACGCTGAGGAGTCCGACCTATGATTTGTGACACACTCAAGTTACAGAGACAC 420
QY 474 GGTGCTCTACCTTCGGGAGCGTGTCTGTGTTGAGGGACCTTGTGTCTTACAGCCA 533
Db 421 GGTGCTCTACCTTCGGGAGCGTGTCTGTGTTGAGGGACCTTGTGTCTTACAGCCA 480
QY 534 GATCCGGGACATGTTCCACCTGGGCTCTTGTGTGACACCGACCTCCAGCGTGC 593
Db 481 GATCCGGGACATGTTCCACCTGGGCTCTTGTGTGACACCGACCTCCAGCGTGC 540
QY 594 TCGAAGAGTCTTCGGGAGCGTGCCTCGAGGAGGAGCTGAGAGATTCGACGACGTA 653
Db 541 TCGAAGAGTCTTCGGGAGCGTGCCTCGAGGAGGAGCTGAGAGATTCGACGACGTA 599
QY 654 CACCACTTCGTGAAAGCCGCTTCGAGAGATTCGCTCGCGCAAAAGATATGCCGA 713
Db 600 CACCACTTCGTGAAAGCCGCTTCGAGAGATTCGCTCGCGCAAAAGATATGCCGA 659
QY 714 TGTGATCATCCAGAGAGTGAACAATATGTTGCCATCAACCTGCTGCGACGACAT 773
Db 660 TGTGATCATCCAGAGAGTGAACAATATGTTGCCATCAACCTGCTGCGACGACAT 719
QY 774 CCAGAGCATTTCTGATGATGTCATCTGCAAAATGACACCGAGAGGTCATGAGGCGAG 833
Db 720 CCAGAGCATTTCTGATGATGTCATCTGCAAAATGACACCGAGAGGTCATGAGGCGAG 779
QY 834 CTACAGAGGACCTTTTCTGAGCCAGGGGACCACTTGGAGTGTCTGCTGGCAACG 893
Db 780 -TACAAAGCGGACCTTTTCTTAGCCAGGGGACCACTTGGAGTGTCTGCTGGCAACG 838
QY 894 GTCAATTTGAGATCCAGAGACAGACCCCACTGAGGGGCTGCGAGGCTCAGAGGCAAGTC 953
Db 839 GTCAATTTGAGATCCAGAGACAGACCCCACTGAGGGGCTGCGAGGCTCAGAGGCAAGTC 898
QY 954 TCCCGCCCGGAGTGTGTTTCAAGGAGCTGA-GCCTGGGGAAGCCCAACCAACCACTGC 1012
Db 899 TCCCGCCCGGAGTGTGTTTCAAGGAGTGTGAGGAGTGTGAGGAGTGTGAGGAGTGTG 958
QY 1013 TTCTCTTCGGGCAACCCAGGGGAGTGTGAGCAGAGGCTTCTCACTCAGAGTGA 1072
Db 959 TTCTCTTCGGGCAACCCAGGGGAGTGTGAGCAGAGGCTTCTCACTCAGAGTGA 1018
QY 1073 AACTCAGATGTGACTCAGACTCAACTGTGCTGGGACACTGACAGGCGTTCTGAGGTTT 1132
Db 1019 AACTCAGATGTGACTCAGACTCAACTGTGCTGGGACACTGACAGGCGTTCTGAGGTTT 1078
```

```
QY 1133 TCAGCCACTTAGGCTGCTGTTGAGGTTAAAGATCCCTAGTCACTGAGAAATGCCACAG 1192
Db 1079 TCAGCCACTTAGGCTGCTGTTGAGGTTAAAGATCCCTAGTCACTGAGAAATGCCACAG 1138
QY 1193 AATGTGAGAGACCTTGGAGGCTTCTGTGAGGAATGTGAGGCAATTAATGGGAAATT 1252
Db 1139 AATGTGAGAGACCTTGGAGGCTTCTGTGAGGAATGTGAGGCAATTAATGGGAAATT 1198
QY 1253 GAGGAGACAGCTTGAACACTGAGGCTGAGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1312
Db 1199 GAGGAGACAGCTTGAACACTGAGGCTGAGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1258
QY 1313 AGAGAGTTTTCAGTCTGATCTGATCTGTTCTTACACAC 1349
Db 1259 AGAGAGTTTTCAGTCTGATCTGATCTGTTCTTACACAC 1295

RESULT 11
AAH04832
ID AAH04832 standard; cDNA; 753 BP.
XX
AC AAH04832;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human cDNA clone (5'-primer) SEQ ID NO:1667.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN EPI074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-0116126.
XX
PR 29-JUL-1999; 99JP-0248036.
PR 27-AUG-1999; 99JP-0300253.
PR 11-JAN-2000; 2000JP-0118776.
PR 02-MAY-2000; 2000JP-0183767.
PR 09-JUN-2000; 2000JP-0241899.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Iehii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
DR WPI; 2001-318749/34.
XX
PT Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs -
XX
PS Claim 1; SEQ ID 1667; 2537pp + CD ROM; English.
XX
CC The present invention describes primer sets for synthesizing 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination
CC of the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesising polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
```

CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.

XX Sequence 753 BP; 158 A; 196 C; 261 G; 135 T; 3 other;

Query Match 38.4%; Score 623; DB 22; Length 753;
Best Local Similarity 100.0%; Pred. No. 2.8e-304;
Matches 623; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 46 GGGGAAAGGGGGGGGGGGGACCCGATGCGCGGAGCCGAGCCGAGATGCTTCGCGG 105
Db 47 GGGGAAAGGGGGGGGGGGGACCCGATGCGCGGAGCCGAGCCGAGATGCTTCGCGG 106
QY 106 GAGGCGAAGACTGCGAGAGCCCGCGCCGAGGCGCAGCCGCGACAGCGGCGCTTC 165
Db 107 GAGGCGAAGACTGCGAGAGCCCGCGCCGAGGCGCAGCCGCGACAGCGGCGCTTC 166
QY 166 CTGATAGGGGTGAGCGCGCGCACTGCGAAGGAGTGCACCTGTGTGAGAGATCATG 225
Db 167 CTGATAGGGGTGAGCGCGCGCACTGCGAAGGAGTGCACCTGTGTGAGAGATCATG 226
QY 226 GAGTGTGCTGGGACAGAGAGAGTGGAAACAGCGGAGCGGAGGAGTGTGATCTTGAGCAG 285
Db 227 GAGTGTGCTGGGACAGAGAGAGTGGAAACAGCGGAGCGGAGGAGTGTGATCTTGAGCAG 286
QY 286 GACAGGTTCTACAAGGTCTGACCGGACAGAGAGCGGCTTGAAAGACAGTAC 345
Db 287 GACAGGTTCTACAAGGTCTGACCGGACAGAGAGCGGCTTGAAAGACAGTAC 346
QY 346 AATTTGACCATCCAGATGCTTTGATATGATTTGATGCAAGAGCTTGAAAGACATC 405
Db 347 AATTTGACCATCCAGATGCTTTGATATGATTTGATGCAAGAGCTTGAAAGACATC 406
QY 406 GTGAGAGGCAAAACGGTGGAGGTGCCGACCTAATTTTGGACACACTCAAGTTACCA 465
Db 407 GTGAGAGGCAAAACGGTGGAGGTGCCGACCTAATTTTGGACACACTCAAGTTACCA 466
QY 466 GAGACACAGGTGTCTACCTGCGAGAGTGTCTGTTTGAAGGCACTTGATTTCTAC 525
Db 467 GAGACACAGGTGTGTCTACCTGCGAGAGTGTCTGTTTGAAGGCACTTGATTTCTAC 526
QY 526 AGCCAGAGATCCGGAATGTTCCACCTGCGCTTTCGTGACACCGACTCCGAGCTC 585
Db 527 AGCCAGAGATCCGGAATGTTCCACCTGCGCTTTCGTGACACCGACTCCGAGCTC 586
QY 586 AGGCTGTCTGGAAGAGTTTCCCGGAGAGTGGCGCGGAGGAGGAGCTTGAGCAGATTCTG 645
Db 587 AGGCTGTCTGGAAGAGTTTCCCGGAGAGTGGCGCGGAGGAGGAGCTTGAGCAGATTCTG 646
QY 646 ACGCAGTACACCACTTCGTGAA 668
Db 647 ACGCAGTACACCACTTCGTGAA 669
```

RESULT 12
AAH75355
ID AAH75355 standard; cDNA; 900 BP.

XX AAH75355;

XX 02-OCT-2001 (first entry)

XX Human uridine kinase encoding cDNA.

XX Human, uridine kinase; UK; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

FT CDS 8..838
FT /<tag= a
FT /product= "uridine kinase"
FT /note= "Claimed in claim 1"

XX CN1287172-A.

XX 14-MAR-2001.

XX 07-SEP-1999; 99CN-0118818.

XX 07-SEP-1999; 99CN-0118818.

XX (YFPU-) UNIV FUDAN.

XX Yu L, Zhao Y, Zhang H;

XX WPI; 2001-409529/44.

XX P-PDB; AAG64506.

XX Human uridine kinase and its coding sequence, preparation and
XX application -

XX Claim 1; Page 15(Disclosure); 20pp; Chinese.

XX The invention relates to human uridine kinase (UK).

XX Sequence 900 BP; 201 A; 237 C; 292 G; 170 T; 0 other;

Query Match 35.2%; Score 572; DB 22; Length 900;
Best Local Similarity 99.6%; Pred. No. 1.7e-278;
Matches 722; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
QY 265 AAGGTGTCACTCTGAGCCAGAGCAGGTTCTACAAAGTCTGACGCGCAGAGCAAGGCC 324
Db 176 AAGGTGTCACTCTGAGCCAGAGCAGGTTCTACAAAGTCTGACGCGCAGAGCAAGGCC 325
QY 325 AAGGCTTGAAGAGACAGTACAAATTTTGAACATCCAGATGCTTGAATATGATTTGATG 384
Db 236 AAGGCTTGAAGAGACAGTACAAATTTTGAACATCCAGATGCTTGAATATGATTTGATG 385
QY 385 CACAGGACTCTGAAGAACATCGTGGAGGCAAAACGGTGGAGGTGCGACCTATGATTTT 444
Db 296 CACAGGACTCTGAAGAACATCGTGGAGGCAAAACGGTGGAGGTGCGACCTATGATTTT 445
QY 445 GTGACACACTCAAGTTTACAGAGACACAGTGTCTACCTCGGACGATGTTCTGTTT 504
Db 356 GTGACACACTCAAGTTTACAGAGACACAGTGTGTCTACCTCGGACGATGTTCTGTTT 504
QY 505 GAGGGCATCTGGGTGTTCTACAGCCAGAGATCCGGAATGTTCCACTGCGGCTTCTT 564
Db 416 GAGGGCATCTGGGTGTTCTACAGCCAGAGATCCGGAATGTTCCACTGCGGCTTCTT 475
QY 416 GAGGGCATCTGGGTGTTCTACAGCCAGAGATCCGGAATGTTCCACTGCGGCTTCTT 475
QY 565 GTGACACCGACTCTCGAGCGTCAAGGCTGTCTCGAAGAGTTCTCCGGGACGTGCGCGAGG 624
Db 476 GTGACACCGACTCTCGAGCGTCAAGGCTGTCTCGAAGAGTTCTCCGGGACGTGCGCGAG 624
QY 625 AGGACCTGAGAGAGATTTCTGACAGCACTACCACTTGTGAAAGCGGCTTTCAGAG 684
Db 536 AGGACCTGAGAGAGATTTCTGACAGCACTACCACTTGTGAAAGCGGCTTTCAGAG 684
QY 685 TTTCGCTGCGGCAAAAGAGTATGCGGATGTGATATCCACAGAGAGTGGCAATATG 744
Db 596 TTTCGCTGCGGCAAAAGAGTATGCGGATGTGATATCCACAGAGAGTGGCAATATG 744
QY 745 GTTGCCATCAACTGATGTCGAGCAGCATCCAGACATTTCTGAATGATGATCTGCAAA 804
Db 656 GTTGCCATCAACTGATGTCGAGCAGCATCCAGACATTTCTGAATGATGATCTGCAAA 804
QY 805 TGGCACCGAGAGAGTTCATGAGCGGAGCTTCAAGCGGACCTTTTCTGAGCCAGGAG 864
Db 716 TGGCACCGAGAGAGTTCATGAGCGGAGCTTCAAGCGGACCTTTTCTGAGCCAGGAG 864
```


PR 07-JUL-2000; 2000US-0216647.
 PR 07-JUL-2000; 2000US-0216880.
 PR 11-JUL-2000; 2000US-0217487.
 PR 11-JUL-2000; 2000US-0217496.
 PR 14-JUL-2000; 2000US-0218290.
 PR 26-JUL-2000; 2000US-0220963.
 PR 26-JUL-2000; 2000US-0220964.
 PR 14-AUG-2000; 2000US-0224518.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0225213.
 PR 14-AUG-2000; 2000US-0225214.
 PR 14-AUG-2000; 2000US-0225266.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225447.
 PR 14-AUG-2000; 2000US-0225757.
 PR 14-AUG-2000; 2000US-0225758.
 PR 14-AUG-2000; 2000US-0225759.
 PR 16-AUG-2000; 2000US-0226279.
 PR 22-AUG-2000; 2000US-0226861.
 PR 22-AUG-2000; 2000US-0226868.
 PR 22-AUG-2000; 2000US-0227182.
 PR 23-AUG-2000; 2000US-0227009.
 PR 30-AUG-2000; 2000US-02289924.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 01-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 05-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 08-SEP-2000; 2000US-0232081.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.

PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX Rosen CA, Barash SC, Ruben SM;
 PI WPI; 2001-451937/48.
 XX P-PSDB; ABB03475.
 DR
 XX
 PT Isolated polypeptide for treating, preventing and/ or prognosing
 PT disorders related to the musculoskeletal system including
 PT musculoskeletal cancers and also for testing and detection e.g.
 PT diagnosis -
 PT
 XX
 PS Claim 1; SEQ ID NO 399; 781bp + Sequence Listing; English.
 XX
 CC The invention relates to novel genes (AAU34669-AAU37666) and proteins
 CC (ABB03087-ABB04109) associated with the musculoskeletal system useful
 CC for preventing, treating or ameliorating medical conditions e.g. by
 CC protein or gene therapy. The genes are isolated from a range of human
 CC tissues disclosed in the specification. The nucleic acids, proteins,
 CC antibodies and (ant)agonists are useful in the diagnosis, treatment

CC and prevention of: (a) cancer, e.g. breast and ovarian cancer and
 CC other cancers of the adrenal gland, bone, bone marrow, breast,
 CC gastrointestinal tract, liver, lung, or urogenital; (b) immune
 CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
 CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
 CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis;
 CC (c) cardiovascular disorders such as myocardial ischaemia; (d) wound
 CC healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy;
 CC and (f) infectious diseases such as viral, bacterial, fungal and
 CC parasitic infections.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 CC
 XX Sequence 222 BP: 44 A; 63 C; 75 G; 38 T; 2 other;
 Query Match 3.6%; Score 58; DB 22; Length 222;
 Best Local Similarity 100.0%; Pred. No. 1.1e-18;
 Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 GY 1521 CCACGAGGAGATGTTTCCCTTGTGCTTAAAGGACAGAGAGCGGGA 1578
 Db 32 CCACGAGGAGATGTTTCCCTTGTGCTTAAAGGACAGAGCGGGA 89
 RESULT 15
 ABX58045
 ID ABX58045 standard; cDNA, 222 BP.
 XX
 AC ABX58045;
 XX
 DT 26-FEB-2003 (first entry)
 XX
 DE cDNA encoding novel human musculoskeletal system antigen #389.
 XX
 KW Gene; ss; musculoskeletal system antigen; cancer; metastasis;
 KW re-vascularisation; thrombosis; arteriosclerosis; mineral content;
 KW cardiovascular condition; wound; injury; burn; angiogenesis; ulcer;
 KW post-operative tissue repair; limb regeneration; neuronal growth;
 KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
 KW AIDS-related complex; chondrocyte growth; bone regeneration;
 KW periodontal regeneration; tissue transport; bone graft; skin aging;
 KW keratinocyte growth; hair loss; melanocyte growth; cell proliferation;
 KW cell growth; organ transplant; cell differentiation; body height;
 KW weight; hair colour; eye colour; skin; percentage of adipose tissue;
 KW pigmentation; cosmetic surgery; metabolism; biorhythm; circadian rhythm;
 KW depression; tendency for violence; pain; reproductive capability;
 KW hormone level; endocrine level; appetite; libido; memory; stress;
 KW storage capability; fat content; lipid content; protein content;
 KW carbohydrate content; vitamin content; cofactor content;
 KW nutritional component.
 XX
 XX Homo sapiens.
 XX
 PN US2002147140-A1.
 PD 10-OCT-2002.
 XX
 PD 17-JAN-2001; 2001US-0764877.
 XX
 PR 31-JAN-2000; 2000US-179065P.
 PR 04-FEB-2000; 2000US-180628P.
 PR 28-JUN-2000; 2000US-214886P.
 PR 07-JUL-2000; 2000US-216647P.
 PR 07-JUL-2000; 2000US-216880P.
 PR 11-JUL-2000; 2000US-217487P.
 PR 11-JUL-2000; 2000US-217496P.
 PR 14-JUL-2000; 2000US-218290P.
 PR 26-JUL-2000; 2000US-220963P.
 PR 26-JUL-2000; 2000US-220964P.
 PR 14-AUG-2000; 2000US-224518P.
 PR 14-AUG-2000; 2000US-224519P.
 PR 14-AUG-2000; 2000US-225267P.

PR 14-AUG-2000; 2000US-225268P.
 PR 14-AUG-2000; 2000US-225270P.
 PR 14-AUG-2000; 2000US-225447P.
 PR 14-AUG-2000; 2000US-225757P.
 PR 14-AUG-2000; 2000US-225758P.
 PR 22-AUG-2000; 2000US-226868P.
 PR 30-AUG-2000; 2000US-228924P.
 PR 01-SEP-2000; 2000US-229287P.
 PR 01-SEP-2000; 2000US-229343P.
 PR 01-SEP-2000; 2000US-229344P.
 PR 01-SEP-2000; 2000US-229345P.
 PR 05-SEP-2000; 2000US-229509P.
 PR 05-SEP-2000; 2000US-229513P.
 PR 08-SEP-2000; 2000US-231413P.
 PR 21-SEP-2000; 2000US-234223P.
 PR 21-SEP-2000; 2000US-234274P.
 PR 25-SEP-2000; 2000US-234997P.
 PR 27-SEP-2000; 2000US-235834P.
 PR 29-SEP-2000; 2000US-236327P.
 PR 29-SEP-2000; 2000US-236327P.
 PR 29-SEP-2000; 2000US-236367P.
 PR 29-SEP-2000; 2000US-236368P.
 PR 29-SEP-2000; 2000US-236369P.
 PR 29-SEP-2000; 2000US-236370P.
 PR 02-OCT-2000; 2000US-236802P.
 PR 02-OCT-2000; 2000US-237037P.
 PR 02-OCT-2000; 2000US-237038P.
 PR 02-OCT-2000; 2000US-237039P.
 PR 02-OCT-2000; 2000US-237040P.
 PR 13-OCT-2000; 2000US-239345P.
 PR 20-OCT-2000; 2000US-240960P.
 PR 20-OCT-2000; 2000US-241785P.
 PR 20-OCT-2000; 2000US-241809P.
 PR 01-NOV-2000; 2000US-244617P.
 PR 17-NOV-2000; 2000US-249299P.
 PR 08-DEC-2000; 2000US-251856P.
 PR 08-DEC-2000; 2000US-251868P.
 PR 08-DEC-2000; 2000US-251869P.
 XX
 PA (ROSE/) ROSEN C A.
 PA (RUBEN/) RUBEN S M.
 PA (BARASH/) BARASH S C.
 XX
 PI Rosen CA, Ruben SM, Barash SC;
 DR WPI: 2003-126199/12.
 PT P-Psdb; ABU12769.
 XX
 PS Isolated nucleic acid molecules encoding musculoskeletal system
 XX associated polypeptides, useful for detecting disorders, e.g. cancer -
 XX Claim 1; SEQ ID NO 399; 321pp; English.
 XX
 CC The invention describes an isolated nucleic acid molecule comprising a
 CC sequence encoding musculoskeletal system associated polypeptides useful
 CC for detecting disorders, e.g., cancer or cancer metastases, in animals
 CC or humans. The nucleic acid; stimulates re-vascularisation of ischaemic
 CC tissues associated with conditions such as thrombosis, arteriosclerosis,
 CC and other cardiovascular conditions; treats wounds due to injuries,
 CC burns, post-operative tissue repair, and ulcers; stimulates angiogenesis
 CC and limb regeneration; stimulates neuronal growth; can treat and prevent
 CC neuronal damage occurring in certain disorders or neurodegenerative
 CC conditions, such as, Alzheimer's disease, Parkinson's disease, and
 CC AIDS-related complex; stimulates chondrocyte growth, thus they can be
 CC used to enhance bone and periodontal regeneration and aid in tissue
 CC transports or bone grafts; prevents skin aging due to sunburn by
 CC stimulating keratinocyte growth; prevents hair loss, since FGF family
 CC members activate hair-forming cells and promotes melanocyte growth;
 CC stimulates growth and differentiation of hematopoietic cells and bone
 CC marrow cells when used in combination with other cytokines; maintains
 CC organs before transplantation or for supporting cell culture of primary
 CC tissues; induces tissue of mesodermal origin to differentiate in early
 CC embryos; increases or decreases the differentiation or proliferation of
 CC embryonic stem cells, besides, haematopoietic lineage; modulates

CC mammalian characteristics, such as, body height, weight, hair colour, eye
CC colour, skin, percentage of adipose tissue, pigmentation, size, and shape
CC (e.g., cosmetic surgery); modulates mammalian metabolism; changes
CC mammal's mental state or physical state by influencing biorhythms,
CC circadian rhythms, depression, tendency for violence, tolerance for pain,
CC reproductive capabilities, hormonal or endocrine levels, appetite,
CC libido, memory, or stress; increases or decreases storage capabilities,
CC fat content, lipid, protein, carbohydrate, vitamins, minerals, cofactors
CC or other nutritional components. This sequence encodes a novel human
CC musculoskeletal system antigen.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from the US patent office at
CC ftp.segdata.uspto.gov/sequence.html?DocID=20020147140.

XX Sequence 222 BP; 44 A; 63 C; 75 G; 38 T; 2 other;

Query Match 3.6%; Score 58; DB 25; Length 222;

Best local Similarity 100.0%; Pred. No. 1.1e-18; Mismatches 0; Gaps 0;

Db 1521 CCACTGAGGAGTGTTCCTTCCCTTGTGCTTAAAGGACAGAGCGAGCGGA 1578
32 CCACTGAGGAGTGTTCCTTCCCTTGTGCTTAAAGGACAGAGCGAGCGGA 89

Search completed: November 25, 2003, 02:31:18
Job time : 500.846 secs

This Page Blank (uspto)

score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

OM nucleic - nucleic search, using sw model

10529.410 Million cell updates/sec

Perfect score: 1624
Sequence: 1 gtgggtcgctcgactc.....ccagctgtgcggggccag 1624

Gapor 60.0 , Gapext 60.0

Word size : 15

99185

Maximum DB seq length: 200000000000

Post-processing: Listing first 45 summaries

Database : GenEmb

- 1: gb_da:*
- 2: gb_hcg:*
- 3: gb_in:*
- 4: gb_om:*
- 5: gb_ov:*
- 6: gb_pat:*
- 7: gb_ph:*
- 8: gb_pl:*
- 9: gb_pr:*
- 10: gb_ro:*
- 11: gb_ses:*
- 12: gb_sy:*
- 13: gb_un:*
- 14: gb_vi:*
- 15: em_ba:*
- 16: em_fun:*
- 17: em_hum:*
- 18: em_in:*
- 19: em_mu:*
- 20: em_or:*
- 21: em_ov:*
- 22: em_ov:*
- 23: em_pat:*
- 24: em_ph:*
- 25: em_pl:*
- 26: em_ro:*
- 27: em_ses:*
- 28: em_un:*
- 29: em_vi:*
- 30: em_hcg_hum:*
- 31: em_hcg_inv:*
- 32: em_hcg_other:*
- 33: em_hcg_mus:*
- 34: em_hcg_pln:*
- 35: em_hcg_rod:*
- 36: em_hcg_mam:*
- 37: em_hcg_vrt:*
- 38: em_sy:*
- 39: em_hcgo_hum:*
- 40: em_hcgo_mus:*
- 41: em_hcgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match Length	ID	Description
1	1624	100.0	1624	6	AX449217 Sequence
2	1304	80.3	2160	6	BD157613 Primer fo
3	1304	80.3	2160	6	AX022317 Homo sapi
4	920	56.7	1022	9	AF237290 Homo sapi
5	880	54.2	147492	9	AL358781 Homo sapi
6	834	51.4	834	6	AX449219 Human DNA
7	783	48.2	834	9	AF254133 Sequence
8	654	40.3	2072	9	BC015547 Homo sapi
9	653	40.2	2228	9	AK057848 Homo sapi
10	620	39.4	1395	9	AF125106 Homo sapi
11	603	38.4	753	6	BD146824 Primer fo
12	338	20.8	734	6	AX540411 Sequence
13	36	2.2	660	11	BV077668 S212P622
14	36	2.2	1610	10	M508UK1
15	36	2.2	1559	10	BC025146
16	36	2.2	23310	2	AC097693 Rattus no
17	36	2.2	241882	10	AL808027 Mouse DNA
18	36	2.2	244105	2	AC098897 Rattus no
19	36	2.2	254577	2	AC068894 Mus muscu
20	33	2.0	254993	2	AC078885 Mus muscu
21	23	1.4	336	9	BT006860 Homo sapi
22	23	1.4	336	12	BT007548 Homo sapi
23	23	1.4	383	11	G41891 SHGC-35183
24	23	1.4	744	9	AB062451 Homo sapi
25	23	1.4	1121	9	AF236637 Homo sapi
26	23	1.4	1209	9	BC002806 Homo sapi
27	23	1.4	1322	6	AI135546 Sequence
28	23	1.4	99866	6	AL358115 Human DNA
29	23	1.4	164246	2	AC016371 Homo sapi
30	23	1.4	178657	2	AC140447 Mus muscu
31	22	1.4	841	11	BV017303 S212P6355
32	22	1.4	952	10	M508UK1A
33	22	1.4	13852	10	AB026490
34	22	1.4	143291	9	HS16369
35	22	1.4	165248	2	AC102308 Human DNA
36	22	1.4	192336	2	AC118474 Mus muscu
37	22	1.4	204529	2	AC021445 Mus muscu
38	22	1.4	211075	10	AC076974 Mus muscu
39	22	1.4	221847	2	AC127929 Rattus no
40	22	1.4	231524	2	AC115181 Rattus no
41	22	1.4	233137	2	AC127758 Rattus no
42	22	1.4	265928	2	AC093937 Rattus no
43	22	1.4	271603	2	AC106318 Rattus no
44	22	1.4	284499	2	AC134366 Rattus no
45	21	1.3	448	10	AF238445 Cricetulus

ALIGNMENTS

RESULT 1				
AX449217				
LOCUS	AX449217	1624 bp	DNA	linear
DEFINITION	Sequence 1 from Patent WO0202761.			
ACCESSION	AX449217			
VERSION	AX449217.1	GI:21697994		
KEYWORDS	.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
REFERENCE	1			
AUTHORS	Glucksmann, M.A.			
TITLE	57658, a human uridine kinase and uses thereof			
JOURNAL	Patent: WO 0202761-A 1 10-JAN-2002;			

FEATURES Millennium Pharmaceuticals, Inc. (US)
Location/Qualifiers
1..1624
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
94..927
/note="unamed protein product"
/codon_start=1
/protein_id="CAD37771.1"
/db_xref="GI:21697995"
/translation="MASAGEDCESPAPEADRPHPRLIGVSGTASGKSTCEIK
ELIGONEVEORORRVILSODRFYKLTABOKAKLGOVNPDPFADLHRLTK
NIVEKTEVPTDPTVTHSLPETHVYVPADVLFEGLIPLVSGOELRDMFHLPLPDT
DSVRLSRVLRVDRRGRDLEQILTYTTTKPAFESECLPTKXIVADVITPRGVDMV
AIIIVHIDILNGDIKMKHGGNSNRSYKRTFSEPDHPMLTSGKSHLSSSRP
H"
BASE COUNT 354 a 427 c 506 g 337 t
ORIGIN
Query Match 100.0%; Score 1624; DB 6; Length 1624;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1624; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GTGGAGTCCGCTCCGACCTGGCGGCGCGCGCGCGCGGGAAGGGGCGGGC 60
DB 1 GTGGAGTCCGCTCCGACCTGGCGGCGCGCGCGCGCGCGGGAAGGGGCGGGC 60
QY 61 GCGGGGACCCGATCGCGGAGCGGAGCGGAGTGGCTTCGCGGGAGGCGAAGCTGC 120
DB 61 GCGGGGACCCGATCGCGGAGCGGAGCGGAGTGGCTTCGCGGGAGGCGAAGCTGC 120
QY 121 GAGAGCCCCGCGCGCGGAGCGGAGCGGAGTGGCTTCGCGGGAGGCGAAGCTGC 180
DB 121 GAGAGCCCCGCGCGCGGAGCGGAGCGGAGTGGCTTCGCGGGAGGCGAAGCTGC 180
QY 181 GCGGCACTGCGGAGCGGAGCGGAGTGGCTTCGCGGGAGGCGAAGCTGC 240
DB 181 GCGGCACTGCGGAGCGGAGCGGAGTGGCTTCGCGGGAGGCGAAGCTGC 240
QY 241 AACGAGGTGAAACGCGGAGCGGAGTGGCTTCGCGGGAGGCGAAGCTGC 300
DB 241 AACGAGGTGAAACGCGGAGCGGAGTGGCTTCGCGGGAGGCGAAGCTGC 300
QY 301 GTCTGACGCGAGCGAAGGCCCAAGCCTTGAAAGACATTTTGAACATCCA 360
DB 301 GTCTGACGCGAGCGAAGGCCCAAGCCTTGAAAGACATTTTGAACATCCA 360
QY 361 GATGCCCTTTGATATGATTTGATGACAGACTCTGAAGACATCTGAGAGGCGAAGC 420
DB 361 GATGCCCTTTGATATGATTTGATGACAGACTCTGAAGACATCTGAGAGGCGAAGC 420
QY 421 GTGAGAGTCCGACCTATGATTTTGTGACACACTCAAGGTTACAGAGACACGGTGTG 480
DB 421 GTGAGAGTCCGACCTATGATTTTGTGACACACTCAAGGTTACAGAGACACGGTGTG 480
QY 481 TACCCGCGGAGCGGAGTCTGTTTGAAGGACCTTGTGATCTTCAACCCAGGAGATCCGG 540
DB 481 TACCCGCGGAGCGGAGTCTGTTTGAAGGACCTTGTGATCTTCAACCCAGGAGATCCGG 540
QY 541 GACATGTTCCACTGCGGCTCTTGTGAGACACGACTCCGACGTCAGGCTGTCTCGAAGA 600
DB 541 GACATGTTCCACTGCGGCTCTTGTGAGACACGACTCCGACGTCAGGCTGTCTCGAAGA 600
QY 601 GTTCTCCGGGACGTGCGCGGAGGAGGACCTTGAGAGAGATTCTGACGACGACACACC 660
DB 601 GTTCTCCGGGACGTGCGCGGAGGAGGACCTTGAGAGAGATTCTGACGACGACACACC 660
QY 661 TTGTGTAAGCGGCTTCGAGAGATTTCTGCTGCGGCAAGAAAGATATGCCGATGTATC 720
DB 661 TTGTGTAAGCGGCTTCGAGAGATTTCTGCTGCGGCAAGAAAGATATGCCGATGTATC 720
QY 721 ATCCCAAGAGAGTGAACAATATGTTGCCATCAACTGATCGTGAAGACATCCAGGAC 780

DB 721 ATCCCAAGAGAGTGAACAATATGTTGCCATCAACTGATCGTGAAGACATCCAGGAC 780
QY 781 ATTCTGAATGTGATCATCTGCAATATGGCAACGAGAGGGTCCAAATGGCGGAGCTCAAG 840
DB 781 ATTCTGAATGTGATCATCTGCAATATGGCAACGAGAGGGTCCAAATGGCGGAGCTCAAG 840
QY 841 CGGACCTTTTCTGAGCGGAGGACCCACTGAGAGTGTCTGAGGAGGAGGAGGAGTACAT 900
DB 841 CGGACCTTTTCTGAGCGGAGGACCCACTGAGAGTGTCTGAGGAGGAGGAGGAGTACAT 900
QY 901 TTGAGATCCAGCAGACGACCCCACTGAGAGGAGTCCGAGGAGTCTGAGGAGGAGTCCG 960
DB 901 TTGAGATCCAGCAGACGACCCCACTGAGAGGAGTCCGAGGAGTCTGAGGAGGAGTCCG 960
QY 961 CGGACATGTGTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1020
DB 961 CGGACATGTGTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1020
QY 1021 GCGGCAACCCGAGGAGGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1080
DB 1021 GCGGCAACCCGAGGAGGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1080
QY 1081 TGTGTACTCAAGCTCAACTTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1140
DB 1081 TGTGTACTCAAGCTCAACTTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1140
QY 1141 TTAGGCTGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1200
DB 1141 TTAGGCTGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1200
QY 1201 GGAAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1260
DB 1201 GGAAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1260
QY 1261 AGCTGAGACACTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1320
DB 1261 AGCTGAGACACTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1320
QY 1321 TTTTCCAGTGTGATCTGTTCTTACACACTCAACATTAAGTTTGTGAGCA 1380
DB 1321 TTTTCCAGTGTGATCTGTTCTTACACACTCAACATTAAGTTTGTGAGCA 1380
QY 1381 AGTACTTCTTTTACATGTTACATGTCATGTTCTGTTCTGTTCTGTTCTGTTCTGTT 1440
DB 1381 AGTACTTCTTTTACATGTTACATGTCATGTTCTGTTCTGTTCTGTTCTGTTCTGTT 1440
QY 1441 CAAGGCTGTGTGCTCAAACTTAATTTATGACCCAGTGTGAGGAGTCAAGCTG 1500
DB 1441 CAAGGCTGTGTGCTCAAACTTAATTTATGACCCAGTGTGAGGAGTCAAGCTG 1500
QY 1501 GCTTACAGGATATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1560
DB 1501 GCTTACAGGATATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1560
QY 1561 GAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1620
DB 1561 GAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1620
QY 1621 CCAG 1624
DB 1621 CCAG 1624
RESULT 2
BD157613
LOCUS BD157613 2160 bp DNA linear PART 17-JAN-2003
DEFINITION Primer for synthesizing full-length cDNA and use thereof.
ACCESSION BD157613
VERSION BD157613.1 GI:27863371
KEYWORDS JP 2002191363-A/12456.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 2160)
Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
Primer for synthesizing full-length cDNA and use thereof
Patent: JP 2002191363-A 12456 09-JUL-2002;
HELIX RESEARCH INSTITUTE
OS Homo sapiens (human)
PN JP 2002191363-A/12456
PD 09-JUL-2002 JP 2000280990
PF 28-JUL-2000 JP 2000280990
PI TOHIO OTA,TAKAO ISOGAI,TERUO NISHIKAWA,KOJI HAYASHI,KAORU
PI SAITO,
PI JUNICHI YAMAMOTO,SHIZUKO ISHII,TOMOYASU SUGIYAMA,AI WAKAMATSU,
PI KEIICHI NAGAI,TERUJII OTSUKI
PC C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/ PC
10, C12P21/02,C12Q1/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof FH Key
Location/Qualifiers
(95)..(925).
FT CDS Location/Qualifiers
1..2160
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 457 a 591 c 671 g 441 t
ORIGIN
Query Match 80.3%; Score 1304; DB 6; Length 2160;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1304; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 46 GGGGAAAGGGCGGCGCGGCGGAGCCCGATGCGCGGAGCGGAGCGGAGTGGCTTCGGCG 105
DB 47 GGGGAAAGGGCGGCGCGGCGGAGCCCGATGCGCGGAGCGGAGCGGAGTGGCTTCGGCG 106
QY 106 GAGAGGGAAGACTGCGAGAGCCCGCGCGGAGCGGAGCGGAGTGGCTTCGGCG 165
DB 107 GAGAGGGAAGACTGCGAGAGCCCGCGCGGAGCGGAGCGGAGTGGCTTCGGCG 166
QY 166 CTGATAGGGGTGAGCGGCGGAGCTGCGAGCGGAGCGGAGTGGCTTCGGCG 225
DB 167 CTGATAGGGGTGAGCGGCGGAGCTGCGAGCGGAGCGGAGTGGCTTCGGCG 226
QY 226 GAGTTCGCGGAGCAGAGCAGTGTGAACAGCGGAGCGGAGTGGCTTCGGCG 285
DB 227 GAGTTCGCGGAGCAGAGCAGTGTGAACAGCGGAGCGGAGTGGCTTCGGCG 286
QY 286 GACAGGTTCTACAGGTTCTCTGACGCGAGCAGAGGCGCAAGGCTTTGAAAGACAGTAC 345
DB 287 GACAGGTTCTACAGGTTCTCTGACGCGAGCAGAGGCGCAAGGCTTTGAAAGACAGTAC 346
QY 346 AATTGGACCATCAATGCTTTGATTAATGATTTGATGACAGGACTCTGAAGAACATC 405
DB 347 AATTGGACCATCAATGCTTTGATTAATGATTTGATGACAGGACTCTGAAGAACATC 406
QY 406 GTGGAGGCGCAAAAGCGTGTGAGTGTGCGGAGCTATGATTTGTGACACACTCAAGATTACCA 465
DB 407 GTGGAGGCGCAAAAGCGTGTGAGTGTGCGGAGCTATGATTTGTGACACACTCAAGATTACCA 466
QY 466 GAGACCAAGGTGTCTACAGGTTCTCTGACGCGAGCGTGTCTGTTGAGGCGCATTTGTTCTAC 525
DB 467 GAGACCAAGGTGTCTACAGGTTCTCTGACGCGAGCGTGTCTGTTGAGGCGCATTTGTTCTAC 526
QY 526 AGCCAGAGAGATCGGAGACATGTTCCACTGCGCTCTGTGTGACACCCAGCTCCGACGTC 585
DB 527 AGCCAGAGAGATCGGAGACATGTTCCACTGCGCTCTGTGTGACACCCAGCTCCGACGTC 586
QY 586 AGGCTGTCTGAGAGATTTCTCCGCGAGCTGTGCGCGGAGGAGGAGCTTGAGACAGATTCTG 645

DB 587 AGGCTGTCTGAGAGATTTCTCCGCGAGCTGTGCGCGGAGGAGGAGCTTGAGACAGATTCTG 646
QY 646 AGCAGGTACACACCTTGTGAGGCGGCGCTTGAGGAGTGTGCTGCGCGCAAGAG 705
DB 647 AGCAGGTACACACCTTGTGAGGCGGCGCTTGAGGAGTGTGCTGCGCGCAAGAG 706
QY 706 TATGCCGATGATCATCTCCACAGAGAGTGTGAATATGTTGCCATCACTGATCTGTG 765
DB 707 TATGCCGATGATCATCTCCACAGAGAGTGTGAATATGTTGCCATCACTGATCTGTG 766
QY 766 CAGCACATCCAGACATTTGTGAATGTGTGACATCTGCAATGCGACCCAGAGGCTCAAT 825
DB 767 CAGCACATCCAGACATTTGTGAATGTGTGACATCTGCAATGCGACCCAGAGGCTCAAT 826
QY 826 GGGCGGAGTACAAAGCGGACCTTTTGTGAGCGAGGAGACCACTGGGATGCTGACCTCT 885
DB 827 GGGCGGAGTACAAAGCGGACCTTTTGTGAGCGAGGAGACCACTGGGATGCTGACCTCT 886
QY 886 GGCMAACGCTCACATTTGAGTGTGACAGACAGACCCCACTGAGGCGCTCCGAGCTCAG 945
DB 887 GGCMAACGCTCACATTTGAGTGTGACAGACAGACCCCACTGAGGCGCTCCGAGCTCAG 946
QY 946 GGCAGGTTCTCCGCGCGGCGGATGTGTGAGGAGTGTGAGGAGTGTGAGGAGTGTGAGG 1005
DB 947 GGCAGGTTCTCCGCGCGGCGGATGTGTGAGGAGTGTGAGGAGTGTGAGGAGTGTGAGG 1006
QY 1006 CCACTGCTTCTCTGCGGCGGCGGAGGAGTGTGAGGAGGAGTGTGAGGAGTGTGAGG 1065
DB 1007 CCACTGCTTCTCTGCGGCGGCGGAGGAGTGTGAGGAGGAGTGTGAGGAGTGTGAGG 1066
QY 1066 GAGTGGAACTCAGATGTGTGACTGAGACTCAACTTGTGAGGAGTGTGAGGAGTGTGAGG 1125
DB 1067 GAGTGGAACTCAGATGTGTGACTGAGACTCAACTTGTGAGGAGTGTGAGGAGTGTGAGG 1126
QY 1126 GAGTGGTTCAGCCTTGAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1185
DB 1127 GAGTGGTTCAGCCTTGAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1186
QY 1186 GGCACAGATGTGAGGAGGAGTGTGAGGAGTGTGAGGAGTGTGAGGAGTGTGAGGAGTGTG 1245
DB 1187 GGCACAGATGTGAGGAGGAGTGTGAGGAGTGTGAGGAGTGTGAGGAGTGTGAGGAGTGTG 1246
QY 1246 GGAATTGAGAGAGAGCTTGAAGAGCTGAGGAGTGTGAGGAGTGTGAGGAGTGTGAGGAG 1305
DB 1247 GGAATTGAGAGAGAGCTTGAAGAGCTGAGGAGTGTGAGGAGTGTGAGGAGTGTGAGGAG 1306
QY 1306 CAGTGGAGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1349
DB 1307 CAGTGGAGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1350
RESULT 3
AK022317 2160 bp mRNA linear pri 01-AUG-2002
LOCUS Homo sapiens CDNA FLJ12255 fis. clone MAMMAL1001476, highly similar
DEFINITION TO URIDINE KINASE (EC 2.7.1.48).
ACCESSION AK022317
VERSION AK022317.1 GI:10433687
KEYWORDS oligo capping; fls (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 Isogai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y.,
Nishikawa,T., Nagai,K., Sugano,S., Shiratori,A., Sudo,H.,
Magatsuma,M., Hosoi,T., Kaku,Y., Kodaira,H., Kondo,H.,
Sugawara,M., Takahashi,M., Chiba,Y., Ishida,S., Murakami,K.,
Ono,Y., Takiguchi,S., Metanabe,S., Kimura,K., Murakami,K.,
Ishii,S., Kawai,Y., Saito,K., Yamamoto,J., Wakamatsu,A.,
Nakamura,Y., Nagahara,K., Masuko,Y., Niinomiya,K. and Iwayanagi,T.
NEDO human cDNA sequencing project
TITLE
JOURNAL
Unpublished

REFERENCE 2 (bases 1 to 2160)
AUTHORS Iwagai,T. and Otsuki,T.
TITLE Direct Submission
JOURNAL Submitted (23-AUG-2000) Takao Iwagai, Helix Research Institute,
Genomics Laboratory, 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
(E-mail:genomic@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of
International Trade and Industry of Japan, cDNA full insert
sequencing: Research Association for Biotechnology, cDNA library
construction, 5'- & 3'-end one pass sequencing and clone selection:
Helix Research Institute (supported by Japan Key Technology Center
etc.) and Department of Virology, Institute of Medical Science,
University of Tokyo.

FEATURES
source Location/Qualifiers
1..2160
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MAMMA1001476"
/tissue_type="Mammary gland"
/clone_id="MAMMA1"
/note="Cloning vector: pME18SFL3"
95..928
/note="unnamed protein product"
/codon_start=1
/protein_id="BAB14010.1"
/db_xref="GI:10433688"
/translation="MASAGEBDCESPAERDRHQRFLIGVSGGTASGKSTCEKIM
ELIGONEVEORRNVILISQDRFYKVLTAQKAKALKQYNFPHPDFAEDMLMRLLK
NIVEGTEVEPTVDVFTHSRLPETTVVYPADVLFEGILVFEQIEIDMDELFLDT
DSVRLSRVLRDVRGRDLLOILTOYTTVKPAFEFRCITPKYADVILPRVDNV
AIVLIVGHIDILNGDICKMHRGSGNSRSTKRTPSEGDHPGLTSGKSHLSSSRP
H"

CDS
BASE COUNT 457 a 591 c 671 g 441 t
ORIGIN

Query Match 80.3%; Score 1304; DB 9; Length 2160;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1304; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 46 GGGGAAAGGGGCGGCGGGGACCCGATGCGCGGAGCGAGCGGAGATGCTTCGGCG 105
Db 47 GGGGAAAGGGGCGGCGGGGACCCGATGCGCGGAGCGGAGCGGAGATGCTTCGGCG 106

Qy 106 GGAAGCGGAAGTCTCGAGAGCCCGCGCGCGGAGCCGATCCGCAACAGCGCCCTTC 165
Db 107 GGAAGCGGAAGTCTCGAGAGCCCGCGCGCGGAGCCGATCCGCAACAGCGCCCTTC 166

Qy 166 CTGATAGGGGTGAGCGGCGGCACTGCGAGCGGGAAGTGCACCGTGTGAGGAAGATCATG 225
Db 167 CTGATAGGGGTGAGCGGCGGCACTGCGAGCGGGAAGTGCACCGTGTGAGGAAGATCATG 226

Qy 226 GAGTTGCTGGGACGAAAGAGGTGAAACAGCGGCGAGCGGAGGTGATCTTACGCGCAG 285
Db 227 GAGTTGCTGGGACGAAAGAGGTGAAACAGCGGCGAGCGGAGGTGATCTTACGCGCAG 286

Qy 286 GACAGGTTCTCAAGAGTCTGACGCGAGAGCAAGAGCCGCTTGAAAAGACAGTAC 345
Db 287 GACAGGTTCTCAAGAGTCTGACGCGAGAGCAAGAGCCGCTTGAAAAGACAGTAC 346

Qy 346 AATTGACCATCGAGATGCTTGTATGATGATTGATGACAGAGACTCTGAAGAATC 405
Db 347 AATTGACCATCGAGATGCTTGTATGATGATTGATGACAGAGACTCTGAAGAATC 406

Qy 406 GTGAGAGGCAAAACGATGAGAGTCCGACCTATGATTTTGTACACTCAAGTTTACCA 465
Db 407 GTGAGAGGCAAAACGATGAGAGTCCGACCTATGATTTTGTACACTCAAGTTTACCA 466

Qy 466 GAGACCAACGATGCTTACCTTCCGAGAGTGGTCTTGTGAGGAGCATCTTGTCTTAC 525
Db 467 GAGACCAACGATGCTTACCTTCCGAGAGTGGTCTTGTGAGGAGCATCTTGTCTTAC 526

Qy 526 AGCCAGAGAGATCGGGACATGTTCCACTGCGCTCTTCTGTGAGACACCACTCCGACGTC 585

Db 527 AGCCAGAGAGATCGGGACATGTTCCACTGCGCTCTTCTGTGAGACACCACTCCGACGTC 586

Qy 586 AGGCTGTCTCGAAGAGTCTCCGGGAGCTGCGCCGAGGAGGAGGACTGTGAGCATTTCTG 645
Db 587 AGGCTGTCTCGAAGAGTCTCCGGGAGCTGCGCCGAGGAGGAGGACTGTGAGCATTTCTG 646

Qy 646 AGCCATTAACACACTTTCGTGAAGCCGGGCTTCGAGAGTTCGCTGCGCAAGAAAG 705
Db 647 AGCCATTAACACACTTTCGTGAAGCCGGGCTTCGAGAGTTCGCTGCGCAAGAAAG 706

Qy 706 TATCGGATGTGATCATCCCAAGAGAGTGAACAATATGTTGCCATCAACTGATCTG 765
Db 707 TATCGGATGTGATCATCCCAAGAGAGTGAACAATATGTTGCCATCAACTGATCTG 766

Qy 766 CAGCAGATCGAGACATTTGTAATGATGATCATCTGCAATGCGACCGAGAGGTTCAAT 825
Db 767 CAGCAGATCGAGACATTTGTAATGATGATCATCTGCAATGCGACCGAGAGGTTCAAT 826

Qy 826 GGGCGGAGCTTCAAGCGGACCTTTTCTGAGCCAGGGGACCACTTGGGATGCTGACCTCT 885
Db 827 GGGCGGAGCTTCAAGCGGACCTTTTCTGAGCCAGGGGACCACTTGGGATGCTGACCTCT 886

Qy 886 GGCAGACGCTCATATTGAGATCCAGCAGACAGACCCCACTGAGGGGCTCCGAGCTCAG 945
Db 887 GGCAGACGCTCATATTGAGATCCAGCAGACAGACCCCACTGAGGGGCTCCGAGCTCAG 946

Qy 946 GGCAGGCTTCCGCGCGGAGTGTGTTGAGGACTGAGCTTGGGAGCGCCACCCAC 1005
Db 947 GGCAGGCTTCCGCGCGGAGTGTGTTGAGGACTGAGCTTGGGAGCGCCACCCAC 1006

Qy 1006 CCACTGCTTCTCTCGGCGGACCCCGAGGAGTGTGAGAGGAGGCTTCTCACTCAG 1065
Db 1007 CCACTGCTTCTCTCGGCGGACCCCGAGGAGTGTGAGAGGAGGCTTCTCACTCAG 1066

Qy 1066 GAGTGGAAACTAGATGTGTCACTAGACTCAACTTGTGAGGACACTGACAGGCTTCT 1125
Db 1067 GAGTGGAAACTAGATGTGTCACTAGACTCAACTTGTGAGGACACTGACAGGCTTCT 1126

Qy 1126 GAGTGTTCAGCCACTTGAAGCTGTGCGGTTTAAAGATCCTCTAGTCACTGAGAAAT 1185
Db 1127 GAGTGTTCAGCCACTTGAAGCTGTGCGGTTTAAAGATCCTCTAGTCACTGAGAAAT 1186

Qy 1186 GCGACAGAAATGCGAGAGAGCCGCGGAGGCTTCTGAGAGAAATGAGGCACTTAATGG 1245
Db 1187 GCGACAGAAATGCGAGAGAGCCGCGGAGGCTTCTGAGAGAAATGAGGCACTTAATGG 1246

Qy 1246 GGAATTTGAGAGACAGCTAGACACTGAGCTGGCTGATGTTTGTGACAGTGAACCA 1305
Db 1247 GGAATTTGAGAGACAGCTAGACACTGAGCTGGCTGATGTTTGTGACAGTGAACCA 1306

Qy 1306 CAGTGGAGAGAGTTTTCAGTCTGATCTGATCTGTTTCAACAC 1349
Db 1307 CAGTGGAGAGAGTTTTCAGTCTGATCTGATCTGATCTGTTTCAACAC 1350

RESULT 4
AF237290 1022 bp mRNA linear PRI 02-AUG-2001
LOCUS AF237290
DEFINITION Homo sapiens uridine-cytidine kinase 1 (UCK1) mRNA, complete cds.
ACCESSION AF237290
VERSION AF237290.1 GI:13506764
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1022)
AUTHORS Van Kompay,A.R., Norda,A., Linden,K., Johansson,M. and Karlsson,A.
TITLE Phosphorylation of uridine and cytidine nucleoside analogs by two
human uridine-cytidine kinases
JOURNAL Mol. Pharmacol. 59 (5), 1181-1186 (2001)
MEDLINE 21203813

PUBMED 11306702
 REFERENCE 2 (bases 1 to 1022)
 AUTHORS Van Rompay, A.R., Lindén, K., Norda, A., Zhu, C., Zheng, X.,
 Johansson, M., and Karlsson, A.
 TITLE Human uridine-cytidine kinase 1 and 2 : rate limiting enzymes
 JOURNAL Unpublished
 REFERENCE 3 (bases 1 to 1022)
 AUTHORS Van Rompay, A.R., Lindén, K., Norda, A., Zhu, C., Zheng, X.,
 Johansson, M., and Karlsson, A.
 TITLE Direct Submission
 JOURNAL Submitted (19-FEB-2000) IMPI, Clinical Virology, Huddinge
 University Hospital, Stockholm 14186, Sweden
 FEATURES
 source
 1..1022
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 1..1022
 /gene="UCK1"
 5..838
 /gene="UCK1"
 /function="phosphorylates uridine and cytidine to uridine
 monophosphate and cytidine monophosphate"
 /codon_start=1
 /product="uridine-cytidine kinase 1"
 /protein_id="AAK28324.1"
 /db_xref="GI:13506765"
 /translation="MASAGDEDESPAPRPHORPFLIGVSGTAGSKVCEKIM
 ELGQNEVEORQRKVVILSDREFYKLTAEOKAKALKGYNFHPAPFNDLMHRTLK
 NIVEGTEVPYDFTVHSRLPETVYVADYVLFEGILVYVQSLRDMFLHLFVDT
 DSPVRLSRRLRDRVRGRDLDEOILTOYTVVFKRPAEPFCLPTKKYADVLIIPRVDMV
 AINLIYQHIQDILINDICRKHRSNGRSYKTRFSEBPHGMLTSGKSHLESSSRP
 H"
 BASE COUNT 227 a 282 c 324 g 189 t
 ORIGIN
 Query Match 56.7%; Score 920; DB 9; Length 1022;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1020; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 481 CACCGACTCGAGAGTCTGTCTGAGAGATTCTCGGGAGCTGCGGGAGGAG 540
 Qy 630 CCGAGAGCAATTCTGACGACAGTACCACTTCCTGGAAGCGGCTTGAGAGTTCTG 689
 Db 541 CCTGAGAGCAATTCTGACGACAGTACCACTTCCTGGAAGCGGCTTGAGAGTTCTG 600
 Qy 690 CCGCGGAGAAAGTATGCGGATGATCATCCAGAGAGTGAATATGTTGCTG 749
 Db 601 CCGCGGAGAAAGTATGCGGATGATCATCCAGAGAGTGAATATGTTGCTG 660
 Qy 750 CATCAACTGATGTCGACACATCCAGAGACATTTGATGATGATCATCTGCAATAGGCA 809
 Db 661 CATCAACTGATGTCGACACATCCAGAGACATTTGATGATGATCATCTGCAATAGGCA 720
 Qy 810 CCGAGAGGGTCCATGAGGCGGAGTCAAGCGGACCTTTCTGAGCCAGGGAGCACCC 869
 Db 721 CCGAGAGGGTCCATGAGGCGGAGTCAAGCGGACCTTTCTGAGCCAGGGAGCACCC 780
 Qy 870 TGGAGATCTGACTCTGCGAAACGGTCAATTGAGTCCAGACAGACCCCACTGAGG 929
 Db 781 TGGAGATCTGACTCTGCGAAACGGTCAATTGAGTCCAGACAGACCCCACTGAGG 840
 Qy 930 GCGTGGAGCTTCAGAGGAGGTTCTCCGCGGATGTTGTTGAGGAGTGAAGCTG 989
 Db 841 GCGTGGAGCTTCAGAGGAGGTTCTCCGCGGATGTTGTTGAGGAGTGAAGCTG 900
 Qy 990 GGAAGCCCAACCAACCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1049
 Db 901 GGAAGCCCAACCAACCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 960
 Qy 1050 GCGCTTCT 1109
 Db 961 GCGCTTCT 1020
 Qy 1110 AC 1111
 Db 1021 AC 1022
 RESULT 5
 AL358781/c
 LOCUS 147492 bp DNA linear PRI 06-OCT-2001
 DEFINITION Human DNA sequence from clone RP11-334J6 on chromosome 9, complete
 sequence.
 ACCESSION AL358781
 VERSION AL358781.19 GI:13751418
 KEYWORDS HTG.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE 1 (bases 1 to 147492)
 AUTHORS Corby, N.
 TITLE Direct Submission
 JOURNAL Submitted (06-OCT-2001) Sanger Centre, Hinxton, Cambridgeshire,
 CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
 COMMENT
 On Apr 21, 2001 this sequence version replaced gi:13396549.
 During sequence assembly data is compared from overlapping clones.
 Where differences are found these are annotated as variations
 together with a note of the overlapping clone name. Note that the
 variation annotation may not be found in the sequence submission
 corresponding to the overlapping clone, as we submit sequences with
 only a small overlap as described above.
 This sequence was finished as follows unless otherwise noted: all
 regions were either double-stranded or sequenced with an alternate
 chemistry or covered by high quality data (i.e., phred quality >= 30);
 an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by at least
 one plasmid subclone or more than one M13 subclone; and the
 assembly was confirmed by restriction digest. The following
 abbreviations are used to associate primary accession numbers given

in the feature table with their source databases: Em, EMBL, SW, SWISSPROT, Tr, TREMBL, Mp, WORMPEP, Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr9> RP11-33406 is from the library RPI1-11.2 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>

FEATURES

location/Qualifiers

1..147492
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="9"
/clone="RP11-33406"
/clone_lib="RPI1-11.2"
misc_feature
30143..30199
/note="Sequence from overlapping clone RP11-643E14 (AU354855). Assembly confirmed by restriction digest."

BASE COUNT 32992 a 36290 c 37577 g 40633 t
ORIGIN

Query Match 54.2%; Score 880; DB 9; Length 147492;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 880; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 745 GTTGCCATCAACCGATCGTGCACATCCAGACATTCGATGCTGATCGTGA 804
Db 129488 GTTGCCATCAACCGATCGTGCACATCCAGACATTCGATGCTGATCG 129429
Qy 805 TGGCACCAGAGAGGCTCAATGCGGAGCTACAAGGAGACCTTTCTGAGCCAGGGAG 864
Db 129428 TGGCACCAGAGAGGCTCAATGCGGAGCTACAAGGAGACCTTTCTGAGCCAGGGAG 129369
Qy 865 CACCTTGAGATGCTGACCTCTGCGAAACGGTCACTTTGAGTCCAGACAGACCCAC 924
Db 129368 CACCTTGAGATGCTGACCTCTGCGAAACGGTCACTTTGAGTCCAGACAGACCCAC 129309
Qy 925 TGAAGGGCTGCCGAGCTCGAGGAGGCTCCGCGCGAGATGCTGAGGAGCTGAG 984
Db 129308 TGAAGGGCTGCCGAGCTCGAGGAGGCTCCGCGCGAGATGCTGAGGAGCTGAG 129249
Qy 985 CTGCGGAGCGCCACCCACACCACTGCTTCTCTCGGCGCACCCAGGGAGGTAGC 1044
Db 129248 CTGCGGAGCGCCACCCACACCACTGCTTCTCTCGGCGCACCCAGGGAGGTAGC 129189
Qy 1045 AGCGAGGCTTCTCTCACTCAGAGTGAATCAATGATGTCACTCAACTCACTTGTCT 1104
Db 129188 AGCGAGGCTTCTCTCACTCAGAGTGAATCAATGATGTCACTCAACTCACTTGTCT 129129
Qy 1105 GGGACATGACAGGCGTTCCTGAGGTTTTCAGCCACTTAGGCTGCTGCTTTAAAGAT 1164
Db 129128 GGGACATGACAGGCGTTCCTGAGGTTTTCAGCCACTTAGGCTGCTGCTTTAAAGAT 129069
Qy 1165 CCCTCTAGTCACTGAGAAATGACAGAAATGTGACAGAAAGCTGAGAGCTTCTGTGAG 1224
Db 129068 CCCTCTAGTCACTGAGAAATGACAGAAATGTGACAGAAAGCTGAGAGCTTCTGTGAG 129009
Qy 1225 GAATGAGGACATTATTTGGGGAATTTGAGAGACAGCTTAGACACTGCTGCTGCTGAT 1284
Db 123008 GAATGAGGACATTATTTGGGGAATTTGAGAGACAGCTTAGACACTGCTGCTGCTGAT 128949
Qy 1285 GTTTTGTGACAGTGAACCAAGTGGAGAGAGTTTTCAGTCTGATCTGCTTTTA 1344
Db 128948 GTTTTGTGACAGTGAACCAAGTGGAGAGAGTTTTCAGTCTGATCTGCTTTTA 128889

Qy 1345 CACACTCACACATAACTCAAAAGTTTGTGAACAAGTACTTCTTTTACATGTTA 1404
Db 128888 CACACTCACACATAACTCAAAAGTTTGTGAACAAGTACTTCTTTTACATGTTA 128829
Qy 1405 CATGCTCATGTTTCTGTTTCTGTTTCAATACAAAGCTGTTGTGCTTACAAAC 1464
Db 128828 CATGCTCATGTTTCTGTTTCTGTTTCAATACAAAGCTGTTGTGCTTACAAAC 128769
Qy 1465 CTATTTTCATGACCGAGTGTGTCAGTCCAGGTCGCTTACACAGATATGGGAGCCAC 1524
Db 128768 CTATTTTCATGACCGAGTGTGTCAGTCCAGGTCGCTTACACAGATATGGGAGCCAC 128709
Qy 1525 TGAGGATGTTTTCCTCCCTTGTGCTTGTGCTTAAAGCAGAGAGCGGATCCCT 1584
Db 128708 TGAGGATGTTTTCCTCCCTTGTGCTTGTGCTTAAAGCAGAGAGCGGATCCCT 128649
Qy 1585 GGAAGCACCCAGCATACACCCAGGCTTGTGCGGGCCAG 1624
Db 128648 GGAAGCACCCAGCATACACCCAGGCTTGTGCGGGCCAG 128609

RESULT 6
AX449219 834 bp DNA linear PAT 03-JUL-2002
LOCUS
DEFINITION
Sequence 3 from Patent WO0202761.
AX449219
ACCESSION
AX449219.1 GI:21697996
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
1
Glucksman, M.A.
57658, a human uridine kinase and uses thereof
Patent: WO 0202761-A 3 10-JAN-2002;
Millennium Pharmaceuticals, Inc. (US)
FEATURES
source
1..834
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

BASE COUNT 194 a 218 c 265 g 157 t
ORIGIN
Query Match 51.4%; Score 834; DB 6; Length 834;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 834; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 94 ATGCTTCGCGGAGGAGGAGAGCTGAGAGCCCGCGGAGGCGGACCGTCCGAC 153
Db 1 ATGCTTCGCGGAGGAGGAGGAGCTGAGAGCCCGCGGAGGCGGACCGTCCGAC 60
Qy 154 CAGCGGCTTCTCTGATGAGGAGGAGGAGCTGAGAGGAGGAGGAGTGAACCTGTGT 213
Db 61 CAGCGGCTTCTCTGATGAGGAGGAGGAGGAGCTGAGAGGAGGAGGAGTGAACCTGTGT 120
Qy 214 GAGAAATCATGAGAGTGTGCTGGGACGAGAGTGAACAGCGGAGCGGAAAGTGTGT 273
Db 121 GAGAAATCATGAGAGTGTGCTGGGACGAGAGTGAACAGCGGAGCGGAAAGTGTGT 180
Qy 274 ATCTGAGCAGAGAGGTTCTTCAAGAGCTCGACGCGGAGAGAGGCAAGGCTTG 333
Db 181 ATCTGAGCAGAGAGGTTCTTCAAGAGCTCGACGCGGAGAGAGGCAAGGCTTG 240
Qy 334 AAGGACAGTATCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 393
Db 241 AAGGACAGTATCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
Qy 394 CTGAAGAAATGTTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 453
Db 301 CTGAAGAAATGTTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 360

QY 454 TCAAGGTTACAGAGACCAAGGTCATGCTACCTGCGGAGTCGTGTTGAGGGGATC 513
Db 361 TCAAGGTTACAGAGACCAAGGTCATGCTACCTGCGGAGTCGTGTTGAGGGGATC 420
QY 514 TTGGTGTTCACAGGAGATCCGGGACATGTTCCAGCTTCCTTCGTGACACC 573
Db 421 TTGGTGTTCACAGGAGATCCGGGACATGTTCCAGCTTCCTTCGTGACACC 480
QY 574 GATCTCCGACGTGAGCTGTCTCCAGAGTTCTCCGGGACGTGCGGAGGAGGACTG 633
Db 481 GACTCCGACGTGAGCTGTCTCCAGAGTTCTCCGGGACGTGCGGAGGAGGACTG 540
QY 634 GAGCAGATTCTGACGAGTACACACCTTCGTGAAGCGGCTTCGAGAGTTCTGCTG 693
Db 541 GAGCAGATTCTGACGAGTACACACCTTCGTGAAGCGGCTTCGAGAGTTCTGCTG 600
QY 694 CCGACAAAGAGTATGCGGATGATCATCCACGAGAGTGAACAATATGTTGCCATC 753
Db 601 CCGACAAAGAGTATGCGGATGATCATCCACGAGAGTGAACAATATGTTGCCATC 660
QY 754 AACCTGATCTGACGACATCCAGGACATTCGATGTTGATGATCTGCAAAATGGCACC 813
Db 661 AACCTGATCTGACGACATCCAGGACATTCGATGTTGATGATCTGCAAAATGGCACC 720
QY 814 GAGGGGTCCAAATGGGCGGAGCTCAAGCGGACCTTTCTGAGCGGAGGACCACTCTGG 873
Db 721 GAGGGGTCCAAATGGGCGGAGCTCAAGCGGACCTTTCTGAGCGGAGGACCACTCTGG 780
QY 874 ATCTGACCTCTGCGCAACCGTCAATTTGGAGTCCAGCAGACCCCACTGA 927
Db 781 ATCTGACCTCTGCGCAACCGTCAATTTGGAGTCCAGCAGACCCCACTGA 834

RESULT 7
LOCUS AF254133 834 bp mRNA linear PRI 02-MAY-2001
DEFINITION Homo sapiens uridine kinase mRNA, complete cds.
ACCESSION AF254133
VERSION AF254133.1 GI:13924749
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Bukarya; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 834)
Ho, Y.S. and Johnson, R.K.
Human uridine kinase from prostate cancer cell line (LNCap)
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 834)
Ho, Y.S. and Johnson, R.K.
TITLE Direct Submission
JOURNAL Submitted (10-APR-2000) Oncology Research, Smlthkline Beecham, 709
Swedeland Road, King of Prussia, PA 19406, USA
FEATURES
source
1. 834
location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1. 834
/note="protein kinase"
/codon_start=1
/product="uridine kinase"
/protein_id="AAK49122.1"
/db_xref="GI:13924750"
translation="MASAGDECDSEPAEDRPHORPFLIGSGTASGKSTVCERIM
ELIGNEVEORORVVILSGDRKYLTAEOKAKLKGQVNFDPHAPAFDNLHRTIK
NIVEGKVEVPTDPTVTHSRLEPTTVVYPADVVLFGILVYSOBIKDMFHLRLFTD
DSPVRLSRVLRGRGRDLQILITQYTTTVKRAPEFCLPTKXIVAVITPRVDNV
AIVLIVQIIDIINGDI CKMRGSGNSRSTKRTFSEPDHGMVLTSGKSHLESSRP
H"

BASE COUNT 194 a 219 c 265 g 156 t
ORIGIN

Query Match 48.2%; Score 783; DB 9; Length 834;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 833; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 94 ATGGCTTCGCGGAGGCGAAGACTGCGAGAGCCCGCGCGGAGGCGGACCGTCCGAC 153
Db 1 ATGGCTTCGCGGAGGCGAAGACTGCGAGAGCCCGCGCGGAGGCGGACCGTCCGAC 60
QY 154 CAGCGGCCCTTCTCTGATATAGGGGTGAGCGCGGCACTGCGAGGGGAAAGTGCACCTGTGT 213
Db 61 CAGCGGCCCTTCTCTGATATAGGGGTGAGCGCGGCACTGCGAGGGGAAAGTGCACCTGTGT 120
QY 214 GAGAAATCATGAGATGCTGCGGAGCAGAACGAGGTGGAACAGCGGAGCGGAAAGTGTG 273
Db 121 GAGAAATCATGAGATGCTGCGGAGCAGAACGAGGTGGAACAGCGGAGCGGAAAGTGTG 180
QY 274 ATCTGAGCCAGAGCAGGTTCTTCAAGAGTCTGACGCGCAGAGCAGAGGCGCAAGGCTTTG 333
Db 181 ATCTGAGCCAGAGCAGGTTCTTCAAGAGTCTGACGCGCAGAGCAGAGGCGCAAGGCTTTG 240
QY 334 AAAGCAGTACAAATTTTGAACATCCAGATGCTTGTGATTAATGATTTGATGACAGACT 393
Db 241 AAAGCAGTACAAATTTTGAACATCCAGATGCTTGTGATTAATGATTTGATGACAGACT 300
QY 394 CTGAAGAATCATGCTGAGGCGCAAAACGCTGAGGTGCGGACCTATGATTTTGTGACACAC 453
Db 301 CTGAAGAATCATGCTGAGGCGCAAAACGCTGAGGTGCGGACCTATGATTTTGTGACACAC 360
QY 454 TCAAGGTTACAGAGACCAAGGTCATCCTTCGCGGACGTGATCTGTTTGAAGGGGATC 513
Db 361 TCAAGGTTACAGAGACCAAGGTCATCCTTCGCGGACGTGATCTGTTTGAAGGGGATC 420
QY 514 TTGGTGTTCACAGGAGATCCGGGACATGTTCCACTGCGCTCTTCGTGACACC 573
Db 421 TTGGTGTTCACAGGAGATCCGGGACATGTTCCACTGCGCTCTTCGTGACACC 480
QY 574 GACTCCGAGTCAAGGTCGTGGAAGATTCCTCGGAGAGTGCAGCGGAGGAGGACTG 633
Db 481 GACTCCGAGTCAAGGTCGTGGAAGATTCCTCGGAGAGTGCAGCGGAGGAGGACTG 540
QY 634 GAGCAGATTCTGACGACATCCAGGACATTCGTTGAAAGCCGCTTCGAGGAGTTCGCTG 693
Db 541 GAGCAGATTCTGACGACATCCAGGACATTCGTTGAAAGCCGCTTCGAGGAGTTCGCTG 600
QY 694 CCGACAAAGAGTATGCGGATGATCATCCACGAGAGTGAACAATATGTTGCCATC 753
Db 601 CCGACAAAGAGTATGCGGATGATCATCCACGAGAGTGAACAATATGTTGCCATC 660
QY 754 AACCTGATCTGACGACATCCAGGACATTCGATGTTGATGATCTGCAAAATGGCACC 813
Db 661 AACCTGATCTGACGACATCCAGGACATTCGATGTTGATGATCTGCAAAATGGCACC 720
QY 814 GAGGGGTCCAAATGGGCGGAGCTCAAGCGGACCTTTCTGAGCGGAGGACCACTCTGG 873
Db 721 GAGGGGTCCAAATGGGCGGAGCTCAAGCGGACCTTTCTGAGCGGAGGACCACTCTGG 780
QY 874 ATCTGACCTCTGCGCAACCGTCAATTTGGAGTCCAGCAGACCCCACTGA 927
Db 781 ATCTGACCTCTGCGCAACCGTCAATTTGGAGTCCAGCAGACCCCACTGA 834

RESULT 8
LOCUS BC015547 2072 bp mRNA linear PRI 04-OCT-2001
DEFINITION Homo sapiens, similar to uridine-cytidine kinase 1, clone MGC:9668
IMAGE:3845821, mRNA, complete cds.
ACCESSION BC015547
VERSION BC015547.1 GI:15930229
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Bukarya; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 2072)
 AUTHORS Strausberg, R.
 TITLE Direct Submission
 JOURNAL Submitted (01-OCT-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgapbs-remail.nih.gov

COMMENT Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: <http://www-shgc.stanford.edu>
 Contact: (Dickson, Mark) mdickpaxil.stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

FEATURES
 source Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAX Plate: 20 Row: f Column: 5
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 13899252.
 Location/Qualifiers
 1..2072
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:9668 IMAGE:3845821"
 /tissue_type="Colon, adenocarcinoma"
 /clone_id="N1H MGC_65"
 /lab_host="DH10B"
 /note="vector: pCMV-SPORT6"
 87..692
 /codon_start=1
 /product="Similar to uridine-cytidine kinase 1"
 /protein_id="AAH15547.1"
 /db_xref="GI:15930230"
 /translation="MASAGGSDCSPPADRPORPPLIGVSGTSGKSTCEKIM ELISQNEVQRKRVILISODRFYKVLAEKAKALGVNPFDPDAFDDLMHRTLK NIVGKTVETPTDFVTHSRLEPTTVVPPADVVFESILVFEYSQELRDMFHLRLFVDT DSDVRLSRDPEVCRCHPTPSGQYGHOPBRAHPSHSEW"

BASE COUNT 455 a 559 c 637 g 421 t

ORIGIN

Query Match 40.3%; Score 654; DB 9; Length 2072;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 654; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

696 GACAAAGAGATATGCGATGATCATCCGACGAGAGTGACAAATGTTGCATCAA 755
 Db 594 GACAAAGAGATATGCGATGATCATCCGACGAGAGTGACAAATGTTGCATCAA 653
 756 CCTGATCGTGACGACATCCAGACATTCTGAATGTCATCTGCAATGGACCGAG 815
 Db 654 CCTGATCGTGACGACATCCAGACATTCTGAATGTCATCTGCAATGGACCGAG 713
 816 AGGCTCAATGGCGGAGCTCAAGCGGACCTTTTCTGAGCCAGGGGACCACTTGAGT 875
 Db 714 AGGCTCAATGGCGGAGCTCAAGCGGACCTTTTCTGAGCCAGGGGACCACTTGAGT 773
 876 GCTGACCTTGGGAAAGGTCATTTGGAGTCCAGAGACGACCCCACTGAGGGGCTGC 935
 Db 774 GCTGACCTTGGGAAAGGTCATTTGGAGTCCAGAGACGACCCCACTGAGGGGCTGC 833
 936 CGAGCCTCAGGGCAGTCTCCCGCCGAGCATGTGTGTCAGAGGACTAGCCTGGGAGCG 995
 Db 834 CGAGCCTCAGGGCAGTCTCCCGCCGAGCATGTGTGTCAGAGGACTAGCCTGGGAGCG 893
 996 CGAGCCTCAGGGCAGTCTCTCTGCGGCGACCCGAGGGAGTGTAGCAGCGAGCCTT 1055

Db 894 CCAACCCACACCACTGCTTCTCTCGGCGACCCACGAGGAGTGTAGCAGAGGCCCTT 953
 Qy 1056 CTTCACTCAGAGATGGAATCTGATGTGATCTGACATCTCACTTCTGGGACACTGAC 1115
 Db 954 CTTCACTCAGAGATGGAATCTGATGTGATCTGACATCTCACTTCTGGGACACTGAC 1013

Qy 1116 AGGCGTCTCAGAGATTTTCAGCCACTTAGGCTGCTGGCTTTAAAGATCCCTAGTGC 1175
 Db 1014 AGGCGTCTCAGAGATTTTCAGCCACTTAGGCTGCTGGCTTTAAAGATCCCTAGTGC 1073

Qy 1176 ACTGAGAAATGCCACAGATGTGAGAGACCTGGAGGCTTTGTGAGAAATGTGAGGC 1235
 Db 1074 ACTGAGAAATGCCACAGATGTGAGAGACCTGGAGGCTTTGTGAGAAATGTGAGGC 1133

Qy 1236 ACATTAATGGGAAATTTAGAGAGACGCCCTAGACACTGAGGCTGCTGATGTTGTGAC 1295
 Db 1134 ACATTAATGGGAAATTTAGAGAGACGCCCTAGACACTGAGGCTGCTGATGTTGTGAC 1193

Qy 1296 AGTGAACCCACAGTGGAGAGAGTCTTTTCCAGTCTGATCTGATCTTACACAC 1349
 Db 1194 AGTGAACCCACAGTGGAGAGAGTCTTTTCCAGTCTGATCTGATCTTACACAC 1247

RESULT 9
 AK057848
 LOCUS
 DEFINITION Homo sapiens cDNA FLJ25119 fis, clone CBR05878, highly similar to URIDINE KINASE (EC 2.7.1.48).
 ACCESSION AK057848
 VERSION AK057848.1 GI:16553809
 KEYWORDS oligo capping; fis (full insert sequence).
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1
 AUTHORS Tashiro, H., Yamazaki, M., Watanabe, K., Kumagai, A., Itakura, S., Fukuzumi, Y., Fujimori, Y., Komiya, M., Suzuki, Y., Hata, H., Nakagawa, K., Mizuno, S., Morinaga, M., Kawamura, M., Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Nishikawa, T., Sugiyama, A., Kawakami, B., Negai, K., Isogai, T., and Sugano, S.
 TITLE NEDO human cDNA sequencing project
 JOURNAL Unpublished
 2 (bases 1 to 2228)
 Submitted (24-OCT-2001) Sumio Sugano, Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure, Human Genome Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan (E-mail: cdna@ims.u-tokyo.ac.jp, Tel: 81-3-5449-5286, Fax: 81-3-5449-5416)
 NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction and 5'-end one pass sequencing: Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure, Human Genome Center; 3'-end one pass sequencing: RAB; clone selection for full insert sequencing: RAB and Helix Research Institute.

FEATURES
 source Location/Qualifiers
 1..2228
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CBR05878"
 /tissue_type="brain"
 /clone_id="CBR"
 /note="cloning vector: pME185FL3"

BASE COUNT 478 a 598 c 696 g 456 t

ORIGIN

Query Match 40.2%; Score 653; DB 9; Length 2228;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 653; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 697 ACAAGAAGTATGCGCATGTATCATCCACGAGAGTGGACAATATGTTGCCATCAAC 756
| | | | |
Db 768 ACAAGAAGTATGCGCATGTATCATCCACGAGAGTGGACAATATGTTGCCATCAAC 827
| | | | |
QY 757 CTGATCGTGCAGCAATCCAGACATTTCTGAATGTGTACATCTGCAATGGACCGAGGA 816
| | | | |
Db 828 CTGATCGTGCAGCAATCCAGACATTTCTGAATGTGTACATCTGCAATGGACCGAGGA 887
| | | | |
QY 817 GGGTCCATGGCGGAGCTACAGCGGACCTTTTCTGAGCCAGGGGACCACTTGGATG 876
| | | | |
Db 888 GGGTCCATGGCGGAGCTACAGCGGACCTTTTCTGAGCCAGGGGACCACTTGGATG 947
| | | | |
QY 877 CTGACCTCTGGCAAGGCTCAATTTGAGTCCAGACGAGACCCCACTGAGGGGCTGCC 936
| | | | |
Db 948 CTGACCTCTGGCAAGGCTCAATTTGAGTCCAGACGAGACCCCACTGAGGGGCTGCC 1007
| | | | |
QY 937 GAGCCTCAGGGCAGGTCTCCGCGCGCATGTGTCTCAGGGACTGAGCCTGGGGAGCC 996
| | | | |
Db 1008 GAGCCTCAGGGCAGGTCTCCGCGCGCATGTGTCTCAGGGACTGAGCCTGGGGAGCC 1067
| | | | |
QY 997 CACCCACCCCACTGCTTCTCTCGCGCACCCCAAGGAGTGTAGCAGCGAGCCTTC 1056
| | | | |
Db 1068 CACCCACCCCACTGCTTCTCTCGCGCACCCCAAGGAGTGTAGCAGCGAGCCTTC 1127
| | | | |
QY 1057 CTGACCTCAGAGTGGAACTCAGATGTGTCACTGACCTGAGCTTGGGACACTGACA 1116
| | | | |
Db 1128 CTGACCTCAGAGTGGAACTCAGATGTGTCACTGACCTGAGCTTGGGACACTGACA 1187
| | | | |
QY 1117 GGGGTTCTGAGGTTTTCAGCACTTAGGCTGTGGGTTTAAAGTCCCTTAGGTTCA 1176
| | | | |
Db 1188 GGGGTTCTGAGGTTTTCAGCACTTAGGCTGTGGGTTTAAAGTCCCTTAGGTTCA 1247
| | | | |
QY 1177 CTGGAATATGCTCAGATGTGCAGAGCCTGGAGGCTTCTGTGAGAAATGTAGGCA 1236
| | | | |
Db 1248 CTGGAATATGCTCAGATGTGCAGAGCCTGGAGGCTTCTGTGAGAAATGTAGGCA 1307
| | | | |
QY 1237 CATTAATGGGGAAATTTAGAGAGACAGCTTGAACACTGGCTGCTGATGTTTGTACA 1296
| | | | |
Db 1308 CATTAATGGGGAAATTTAGAGAGACAGCTTGAACACTGGCTGCTGATGTTTGTACA 1367
| | | | |
QY 1297 GTGAACCCACAGTGGAGAGGTTTTCAGTCTGATCTGATCTTACACAC 1349
| | | | |
Db 1368 GTGAACCCACAGTGGAGAGGTTTTCAGTCTGATCTGATCTTACACAC 1420
| | | | |
RESULT 10
AF125106 1395 bp mRNA linear PRI 07-FEB-2002
LOCUS Homo sapiens uridine kinase mRNA, complete cds.
DEFINITION AF125106
ACCESSION AF125106.1 GI:18568108
VERSION
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Xin,Y.R., Yu,L. and Zhao,S.Y.
TITLE Cloning of a new human cDNA similar to Mus musculus uridine kinase
mRNA
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1395)
AUTHORS Ding,J.B., Yu,L. and Zhao,S.Y.
TITLE Direct Submission
JOURNAL Submitted (02-FEB-1999) Lab of Human Gene Research, Institute of
Genetics, Pudan University, No. 220 Handan Rd., Shanghai 200433,
People's Republic of China
FEATURES
Source
1..1395
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"

CDS
8..838
/codon_start=1
/product="uridine kinase"
/protein_id="AA175943.1"
/db_xref="GI:18568109"
/translation="MASAGEGARARAGANRPHORPDLIGVSGTASGKSTYCKIME
LIGONEVEORHGKVVILSODRFYKVLTAQKAKALAGQNFDPHDAFDDMLMRITKN
IYEGKVEVEPTDFTVHSLPFTTVVYPADVVLFEGILVFSQEIRDMFLRLFPVTD
SDVLSRRLVRVRGRDLEFQILTQYTTTVKRAFEFFCIPTKYAVVILPRVDNVA
INLVHIDILNGLCKMHRGGSNRTYKRTFSEPDHPIMLTSKRSHLESSRPH
"
BASE COUNT 325 a 352 c 411 g 307 t
ORIGIN
Query Match 39.4%; Score 640; DB 9; Length 1395;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 910; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
QY 265 AAGGTGATCCTGAGCCAGACAGAGTCTCAAGAGTCTGACGAGCAGAGAGCC 324
| | | | |
Db 176 AAGGTGATCCTGAGCCAGACAGAGTCTCAAGAGTCTGACGAGCAGAGAGCC 235
| | | | |
QY 325 AAGGCTTGAAGAGACAGTAAATTTTGAACCATCCAGATGCTTTGATATGATTGATG 384
| | | | |
Db 236 AAGGCTTGAAGAGACAGTAAATTTTGAACCATCCAGATGCTTTGATATGATTGATG 295
| | | | |
QY 385 CACAGACTCTGAAGAACATCTGAGAGGCAAAACGCTGAGAGTCCGACTTATGTTT 444
| | | | |
Db 296 CACAGACTCTGAAGAACATCTGAGAGGCAAAACGCTGAGAGTCCGACTTATGTTT 355
| | | | |
QY 445 GTGACACTCAAGGTTTACAGAGACACAGTGTCTACCTCGGACGTTGTTT 504
| | | | |
Db 356 GTGACACTCAAGGTTTACAGAGACACAGTGTCTACCTCGGACGTTGTTT 415
| | | | |
QY 505 GAGGGCATCTTGTGTCTTACACGCGAGAGATCCGGAGATGTTCCACTCGCTTTC 564
| | | | |
Db 416 GAGGGCATCTTGTGTCTTACACGCGAGAGATCCGGAGATGTTCCACTCGCTTTC 475
| | | | |
QY 565 GTGGAACCGACTCTCGACGTCAGGCTGTCTCGAAGGTTCTCCGGACGTGCCCGAGGG 624
| | | | |
Db 476 GTGGAACCGACTCTCGACGTCAGGCTGTCTCGAAGGTTCTCCGGACGTGCCCGAGGG 535
| | | | |
QY 625 AAGGACCTGGAGAGATTTCTGACAGTACACCACTTGTGTAAGCGGCTTCGAGAG 684
| | | | |
Db 536 AAGGACCTGGAGAGATTTCTGACAGTACACCACTTGTGTAAGCGGCTTCGAGAG 595
| | | | |
QY 685 TTCTGCTGCCGCAAGAAAGTATGCGATGTGATCATCCAGAGAGTGAATATG 744
| | | | |
Db 596 TTCTGCTGCCGCAAGAAAGTATGCGATGTGATCATCCAGAGAGTGAATATG 655
| | | | |
QY 745 GTTGCATCAACTGATCTGTGACGACATTCAGAGACATTTGATGTGATCTGCACA 804
| | | | |
Db 656 GTTGCATCAACTGATCTGTGACGACATTCAGAGACATTTGATGTGATCTGCACA 715
| | | | |
QY 805 TGGCAACGAGAGAGGTTCATATGGCGGAGCTCAACAGGACCTTTCTGACCAAGGGAG 864
| | | | |
Db 716 TGGCAACGAGAGAGGTTCATATGGCGGAGCTCAACAGGACCTTTCTGACCAAGGGAG 775
| | | | |
QY 865 CACCTGGAGTGTGACTCTGTGCAACAGGTCAATTTGAGAGTCCAGACAGACCCAC 924
| | | | |
Db 776 CACCTGGAGTGTGACTCTGTGCAACAGGTCAATTTGAGAGTCCAGACAGACCCAC 835
| | | | |
QY 925 TGAAGGCTGCGAGGCTTCAAGGACAGTCTCCGCGCGCATGTGTTCAGGAGCTGAG 984
| | | | |
Db 836 TGAAGGCTGCGAGGCTTCAAGGACAGTCTCCGCGCGCATGTGTTCAGGAGCTGAG 895
| | | | |
QY 985 CCGGGAGAGCCACCAACCAACCACTGCTTCTGCGGAGACCCCAAGGGAGTGTAGC 1044
| | | | |
Db 896 CTTGGAGAGCCACCAACCAACCACTGCTTCTGCGGAGACCCCAAGGGAGTGTAGC 955
| | | | |
QY 1045 AGCAGAGCTTCTCTCACTGAGAGTGAATCTCAGATGTGTCACTCAGACTCAACTTGTCT 1104
| | | | |
Db 956 AGCAGAG-CTTCTCTCACTCAGAGTGAATCTCAGATGTGTCACTCAGACTCAACTTGTCT 1014
| | | | |

Qy	1105	GGGACACTGACAGGGGGTTCCTGAGGTTTCAGCCACTTAGGCTGCTGCGTTAAAGAT	1164
Db	1015	GGGACACTGACAGGGGGTTCCTGAGGTTTCAGCCACTTAGGCTGCTGCGTTAAAGAT	1074
Qy	1165	CCCTCTAGGTCACT	1178
Db	1075	CCCTCTAGGTCACT	1088
RESULT 11			
BD146824		753 bp	DNA
LOCUS			linear
DEFINITION			PAT 17-JAN-2003
ACCESSION			
BD146824.1			GI:27852582
VERSION			JP 2002191363-A/1667.
KEYWORDS			Homo sapiens (human)
SOURCE			Homo sapiens
ORGANISM			Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE			1 (bases 1 to 753)
AUTHORS			Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J., Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
TITLE			Primer for synthesizing full-length cDNA and use thereof
JOURNAL			Patent: JP 2002191363-A 1667 09-JUL-2002; HELIX RESEARCH INSTITUTE
COMMENT			OS Homo sapiens (human) PN JP 2002191363-A/1667 PD 09-JUL-2002 PI 28-JUL-2000 JP 2000280990 PI TOSHIO OTA,TAKAO ISOgai,TETSUO NISHIKAWA,KOJI HAYASHI,KAORU PI SAITO, PI JUNICHI YAMAMOTO,SHIZUKO ISHII,TOMOYASU SUGIYAMA,AI WAKAMATSU, PI KEIICHI NAGAI,TETSUJI OTSUKI PC C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/ PC 10, PC C12P21/02,C12O1/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC Primer for synthesizing full-length cDNA and use thereof FH Key Location/Qualifiers FT source 1..753 FT Location/Qualifiers Location/Qualifiers 1..753 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" 158 a 196 c 261 g 135 t 3 others
FEATURES			
source			
BASE COUNT			
ORIGIN			
Query Match			
Best Local Similarity			38.4%; Score 623; DB 6; Length 753;
Matches			623; Conservative 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;
Qy	46	GGGAGAGGGGCGGGCGCGGGAGACCCGATGCGCGGGAGCGGAGCGCGAGATGGCTTCGCGC	105
Db	47	GGGAGAGGGGCGGGCGCGGGAGACCCGATGCGCGGGAGCGGAGCGCGAGATGGCTTCGCGC	106
Qy	106	GGAGCGGAAGACTGCGAGAGCCCGCGCGCGGAGGCCGATCGCTCCGACACAGCGGCCCTTC	165
Db	107	GGAGCGGAAGACTGCGAGAGCCCGCGCGCGGAGGCCGATCGCTCCGACACAGCGGCCCTTC	166
Qy	166	CTGATAGGGGGGAGAGCGCGGCACTGCGACGGGGGAAGTCGACCGGTGTGAGAAATCATG	225
Db	167	CTGATAGGGGGGAGAGCGCGGCACTGCGACGGGGGAAGTCGACCGGTGTGAGAAATCATG	226
Qy	226	GAGTTGCTCGGACAGAAAGAGGTGGAAACAGCGGACCGGAGCGGAGGTGTCATCTTAGCCAG	285
Db	227	GAGTTGCTCGGACAGAAAGAGGTGGAAACAGCGGACCGGAGGTGTCATCTTAGCCAG	286
Qy	286	GACAGGTTCTACAGGTTCTTGAACGACAGAGCGAAGAGCCAAAGCCCTTGAAGACAGTAC	345

Db	287	GACAGGTTCTACAGAGTCTCTGACGGCAGACGAAAGCCAGGCGCTTGAAAGACAGTAC	346
Qy	346	AATTTTGACATCCAGATGCGCTTTGATATATGATTTGATGACACAGGACTCTGAAACAATC	405
Db	347	AATTTTGACATCTCCAGAGTCCCTTGATATGATTTGATGACACAGGACTCTGAAACAATC	406
Qy	406	GTGAGGGGCAAAAACGGTGGAGGTGCGACCTATGATTTTGTGACACACTCCAAAGTTTACCA	465
Db	407	GTGAGGGGCAAAAACGGTGGAGGTGCGACCTATGATTTTGTGACACACTCCAAAGTTTACCA	466
Qy	466	GAGACACAGGTGTCTACCTTGCGGACGTGTGTTTGTGAGGGCATTTGTGTCTTAC	525
Db	467	GAGACACAGGTGTCTACCTTGCGGACGTGTGTTTGTGAGGGCATTTGTGTCTTAC	526
Qy	526	AGCCAGGAGATCCCGGACATATTTCCACTTGCGGCGCTTCCGTGAGAACCGACTCCGACGTC	585
Db	527	AGCCAGGAGATCCCGGACATATTTCCACTTGCGGCGCTTCCGTGAGAACCGACTCCGACGTC	586
Qy	586	AGGCTGTCTCGAAGATTCCTCGGACGTGCGCCGACGAGGAGGAGACTTGAGACAGATTCGTG	645
Db	587	AGGCTGTCTCGAAGATTCCTCGGACGTGCGCCGACGAGGAGGAGACTTGAGACAGATTCGTG	646
Qy	646	ACGCAGTACACCACTTGTGTGA	668
Db	647	ACGCAGTACACCACTTGTGTGA	669
RESULT 12			
AX540411	AX540411	734 bp	DNA
LOCUS	Sequence 23 from Patent WO02055738.	linear	PAT 23-NOV-2002
DEFINITION	AX540411		
ACCESSION	AX540411.1	GI:25272437	
VERSION			
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Bukacinski, Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	1		
TITLE	Palzer, S.R., Lincoln, S.E., Altus, C.M., Dufour, G.E., Hillman, J.L., Jones, A.L., Dam, T.C., Liu, T.F., Harris, B., Flores, V., Daffo, A., Mawhara, R., Chen, A.J., Chang, S.C., Gerstein, E.H., Peralta, C.H., David, M.H. and Lewis, S.A. Molecules for disease detection and treatment		
JOURNAL	Patent: WO 02055738-A 23 18-JUL-2002; Incyte Genomics, Inc. (US)		
FEATURES.	Location/Qualifiers		
source	1..734		
	/organism="Homo sapiens"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:9606"		
	/note="Incyte ID No: LI:235557.12:2001JUN12"		
BASE COUNT	166 a 181 c 226 g 154 t		7 others
ORIGIN			
Query Match	20.8%	Score 338;	DB 6;
Best Local Similarity	100.0%;	Pred. No. 2.6e-165;	
Matches 338;	Conservative 0;	Mismatches 0;	Indels 0;
			Gaps 0;
Qy	359	CAGATGCGCTTTGATATGATTTGATGACAGAGCTCTGAAAGACATCGTGGAGGGCAAAA	418
Db	12	CAGATGCGCTTTGATATGATTTGATGACAGAGCTCTGAAAGACATCGTGGAGGGCAAAA	71
Qy	419	CGGTGGAGGTGCCGACCTATGATTTTGTGACACACTCAAGTTACAGAGACACGGTGG	478
Db	72	CGGTGGAGGTGCCGACCTATGATTTTGTGACACACTCAAGTTACAGAGACACGGTGG	131
Qy	479	TCTACCGTGGAGGTGGTGTCTGTTTGGAGGCACTTGTTGTTTACACGACGAGGATCC	538
Db	132	TCTACCGTGGAGGTGGTGTCTGTTTGGAGGCACTTGTTGTTTACACGACGAGGATCC	191
Qy	539	GGGACATGTTCCACTGTGGCGCTCTTGTGTGACACCGACTCCGACGTCAAGGTGTCTCGAA	598

Db 192 GGAGCATGTTCCAGCTGCGCCTCTTCTGTGACACCGACTCCGACGTGAGGCTGTCTGAA 251
Qy 599 GAGTTCCTCCGGGACGTGCGCGCGGAGGAGGACCTTGAGGAGATTCTGACCGACATACCA 658
Db 252 GAGTTCCTCCGGGACGTGCGCGCGGAGGAGGACCTTGAGGAGATTCTGACCGACATACCA 311
Qy 659 CTTTCGAGACCGGCTCTGAGGAGATTCTGCGCTGCGG 696
Db 312 CTTTCGAGACCGGCTCTGAGGAGATTCTGCGCTGCGG 349

RESULT 13
BV077668/c
LOCUS 660 bp DNA linear STS 31-MAY-2003
DEFINITION S2126227FDB.T0 CZECHII/E1 Mus musculus STS genomic, sequence
TAGGED SITE.
BV077668
BV077668.1 GI:31193463
STX.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 660)
Wade,C.M., Kulbokas,E.J. III, Kirby,A.W., Zody,M.C., Mullikin,J.C.,
Lander,E.S., Lindblad-Toh,K. and Daly,M.J.
The mosaic structure of variation in the laboratory mouse genome
Nature 420 (6915), 574-578 (2002)
22354684
12466852
COMMENT

TITLE Contact: Kerstin Lindblad-Toh
JOURNAL Whitehead Institute for Biomedical Research, Center for Genome
MEDLINE Research
PUBMED 320 Charles Street, Cambridge, MA 02141, USA
Tel: 6172521477
Fax: 6172580903
Email: kersti@genome.wi.mit.edu
Primer A: None
Primer B: None
STS size: 660
Protocol:
MGS-discovery: Paired-end low-coverage whole genome shotgun reads
were generated from 129S1/SvJmJ, C3H/HeJ, and BALB/cByJ. The MGS
reads were placed uniquely on the MGS33 C57BL/6J assembly and SNP
detection was carried out by SSAHA-SNP. 225,000 reads were
annotated
as STSs and 81,000 SNPs were annotated with alleles from C57BL/6J
and the strain from which the particular read came. The validation
rate for these SNPs was estimated at approximately 98%.

FEATURES
source
1..660
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="CZECHII/E1"
/db_xref="taxon:10090"
/map="2.22-615 32611326-32611918"
/clone_lib="CZECHII/E1"
<1..>660
174 a 196 c 174 g 116 t

STS
BASE COUNT 174 a 196 c 174 g 116 t
ORIGIN

Query Match 2.2%; Score 36; DB 11; Length 660;
Best Local Similarity 100.0%; Pred. No. 6.9e-07;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 528 CCAGAGATCCGGGACATGTTCCACCTGCGCCTCTT 563
Db 435 CCAGAGATCCGGGACATGTTCCACCTGCGCCTCTT 400

RESULT 14
MUSURKI

LOCUS MUSURKI 1810 bp mRNA linear ROD 27-MAR-1997
DEFINITION Mus musculus uridine kinase mRNA, partial cds.
L31783
L31783.1 GI:471980
VERSION
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1810)
Traut,T.W.
The functions and consensus motifs of nine types of peptide
segments that form different types of nucleotide-binding sites
Eur. J. Biochem. 222 (1), 9-19 (1994)
94259063
MEDLINE 8200357
PUBMED 2 (bases 1 to 1810)
REFERENCE Ropp,P.A. and Traut,T.W.
AUTHORS Cloning and expression of a cDNA encoding uridine kinase from mouse
TITLE Arch. Biochem. Biophys. 336 (1), 105-112 (1996)
JOURNAL Arch. Biochem. Biophys. 336 (1), 105-112 (1996)
MEDLINE 97108719
PUBMED 8951040

FEATURES
source
Location/Qualifiers
1..1810
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/tissue_type="Brain"
<1..783
/EC number="2.7.1.48"
/standard_name="uridine-cytidine kinase"
/codon_start=1
/product="uridine kinase"
/protein_id="AAB50568.1"
/db_xref="GI:471981"
/translation="RPPRPFLIGVSGTASGKSTVCKEIMELGONEVDRRORKLVI
LSODCFYKVLTAOKAKALGOYNPHDPAPNDLMHKTLMNIVEGKTVFPTDVT
HRLRPETTYVPADVYLFEGILVPTQERDMFHLRLPDTSDVRLSRVLRDVGR
RDLQILTYIAVFAFEFELCTPKTADAVIIPKVDMMVAINLIVQHIQDILNSDL
CRNRGGPNRNMKRTFPPGDPGVLATGKSHLESSSRPH"
join(37..63,412..426,451..465)
/bound_molecy="ATP at catalytic site"
misc_binding
polya_signal 1790..1795
BASE COUNT 411 a 482 c 512 g 405 t
ORIGIN

Query Match 2.2%; Score 36; DB 10; Length 1810;
Best Local Similarity 100.0%; Pred. No. 6.4e-07;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 528 CCAGAGATCCGGGACATGTTCCACCTGCGCCTCTT 563
Db 384 CCAGAGATCCGGGACATGTTCCACCTGCGCCTCTT 419

RESULT 15
BC025146
LOCUS 1959 bp mRNA linear ROD 16-APR-2003
DEFINITION Mus musculus uridine monophosphate kinase, mRNA (cDNA clone
MGC:36231 IMAGE:4913412), complete cds.
BC025146
BC025146.1 GI:19263563
VERSION
KEYWORDS
SOURCE MGC.
ORGANISM Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1959)
Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

REMARK
COMMENT

Datchenko, L., Marusina, K., Farmer, A. A., Rubin, G. M., Hong, L., Scapleton, M., Soares, M. B., Bonaldo, M. F., Casavant, T. L., Scheetz, T. E., Brownstein, M. J., Ueding, T. B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S. S., Loguigliano, N. A., Peters, G. J., Abramson, R. D., Mullany, S. J., Bosak, S. A., McEwan, P. J., McKernan, K. J., Malek, U. A., Gunaratne, P. H., Richards, S., Worley, K. C., Hale, S., Garcia, A. M., Gay, L. J., Hui, S. W., Villalón, D. K., Muzny, D. M., Sodergren, E. J., Lu, X., Gibbs, R. A., Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A. C., Shcherchenko, Y., Bouffard, G. G., Blakesley, R. W., Touchman, J. W., Green, E. D., Dickson, M. C., Rodriguez, A. C., Grimwood, J., Schmutz, J., Myers, R. M., Butterfield, J. S., Krzywinski, M. I., Skalska, U., Smalins, D. E., Scherch, A., Schein, J. E., Jones, S. J., and Marra, M. A. Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

2 (bases 1 to 1959)
12477932
Strausberg, R.
Direct Submission
Submitted (05-MAR-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: amg@bcm.tmc.edu
Gunaratne, P. H., Garcia, A. M., Lu, X., Hui, S. W., Louised, H., Kowis, C. R., Sneed, A. J., Martin, R. G., Muzny, D. M., Navavati, A. N., Gibbs, R. A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at: <http://image.llnl.gov>
Series: IRAX Plate: 61 Row: P Column: 19
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, Similarity but not identity to protein.

FEATURES
SOURCE

gene
CDS

1. 1959
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="MGC:36231 IMAGE:4913412"
/tissue_type="Salivary gland, 10 week old female mouse"
/clone_id="NCI_CGAP_SG2"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
1. 1959
/gene="Umpk"
/db_xref="LocusID:22245"
/db_xref="MGI:98904"
57. 890
/codon_start=1
/product="Umpk protein"
/protein_id="AAH25146.1"
/db_xref="GI:19263564"
/db_xref="LocusID:22245"
/translation="MASAGGSGSESAPPEARPPFLIGVSGTASGSKTCEKIM ELIGONEVDRORRLVLSODCFKVLTAOKAKLKGYNFDPDAPFNDLMHKTLK NIVEGKVEVPTDFTVTHPEFTTVYPADVLFEGIIIVFOEIRDMFHLPLPDT DSPVRLSRVLRDVQGRDLEQITQYTAFKPAFEFCPLTKIADVITPRGVNMY AINLIYOHIDILNGDLCKRHGGNGRNHRTFPEPGHPVLATGKRSHLESSRP

BASE COUNT H
ORIGIN 466 a 511 c 561 g 421 t
Query Match 2.2%; Score 36; DB 10; Length 1959;
Best Local Similarity 100.0%; Pred. No. 6.4e-07;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 528 CCAGAGATCCGGAGACATGTTCCACCTGCGCCTCTT 563
Db 491 CCAGAGATCCGGAGACATGTTCCACCTGCGCCTCTT 526

Search completed: November 25, 2003, 05:12:17
Job time: 6:32:18 secs


```

QY 61 CAGCGCCCTTCTGATAGGGGTGAGCGGCGGCACTGCGACGGGAAAGTCGCGTGT 120
D 155 CAGCGCCCTTCTGATAGGGGTGAGCGGCGGCACTGCGACGGGAAAGTCGCGTGT 214
QY 121 GAGAAATCATGAGTGTCTGGGACAGAAAGAGTGAACAGCGGACGGGAAAGTGT 180
D 215 GAGAAATCATGAGTGTCTGGGACAGAAAGAGTGAACAGCGGACGGGAAAGTGT 274
QY 181 ATCTGAGCCAGGACAGGTTCTACAGAGTCTGACGGGACAGAAAGGCGGCTTG 240
D 275 ATCTGAGCCAGGACAGGTTCTACAGAGTCTGACGGGACAGAAAGGCGGCTTG 334
QY 241 AAAGGACAGTACATTTTGAACATCCAGATGCTTGTATATGATTTGATGACAGGACT 300
D 335 AAAGGACAGTACATTTTGAACATCCAGATGCTTGTATATGATTTGATGACAGGACT 394
QY 301 CTGAAGAACATCTGAGAGGCGAAACGATGAGGTGCGACCTATGATTTTGTGACAC 360
D 395 CTGAAGAACATCTGAGAGGCGAAACGATGAGGTGCGACCTATGATTTTGTGACAC 454
QY 361 TCAAGGTTACAGAGACCGGTGTCTACCTTGGGACGTGTCTTGTGAGGCGATC 420
D 455 TCAAGGTTACAGAGACCGGTGTCTACCTTGGGACGTGTCTTGTGAGGCGATC 514
QY 421 TTGGTGTCTACAGGACGAGATCCGGGACATGTTCCAGCTGCGCTCTTGTGAGGCGATC 480
D 515 TTGGTGTCTACAGGACGAGATCCGGGACATGTTCCAGCTGCGCTCTTGTGAGGCGATC 574
QY 481 GACTCCGACGTCAGGCTGTCTGAAAGTTCGCGGACGTGCGCGAGGAGGAGCTG 540
D 575 GACTCCGACGTCAGGCTGTCTGAAAGTTCGCGGACGTGCGCGAGGAGGAGCTG 634
QY 541 GAGAGATTCAGAGGACGATACCACTTGTGAAAGCGGCTTGTGAGGAGTTCGCTG 600
D 635 GAGAGATTCAGAGGACGATACCACTTGTGAAAGCGGCTTGTGAGGAGTTCGCTG 694
QY 601 CCGCAAGAAAGTATGCGGATGTATCATCCACAGAGAGTGAACAATATGTTGCCATC 660
D 695 CCGCAAGAAAGTATGCGGATGTATCATCCACAGAGAGTGAACAATATGTTGCCATC 754
QY 661 AACCTGATGTCAGACATCCAGAGACATTCGATGATGATCATCTGCAAAATGCGAC 720
D 755 AACCTGATGTCAGACATCCAGAGACATTCGATGATGATCATCTGCAAAATGCGAC 814
QY 721 GAGAGGTCATGAGGCGGAGCTTCAAGGAGACCTTTTGTGAGGCGGAGACCTTGG 780
D 815 GAGAGGTCATGAGGCGGAGCTTCAAGGAGACCTTTTGTGAGGCGGAGACCTTGG 874
QY 781 ATGCTGACCTCTGCAAAAGGTCATTTGAGATCCAGAGACAGACCCCACTGA 834
D 875 ATGCTGACCTCTGCAAAAGGTCATTTGAGATCCAGAGACAGACCCCACTGA 928

```

RESULT 2

```

US-09-536-647-1
; Sequence 1, Application US/09536647
; Patent No. 6579708
; GENERAL INFORMATION:
; APPLICANT: Ho, Yen Sen
; APPLICANT: Johnson, Randall
; TITLE OF INVENTION: No. 6579708el Human Uridine Kinase
; FILE REFERENCE: G550020
; CURRENT APPLICATION NUMBER: US/09/536,647A
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 834
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(780)

```

the SIDI =

US-09-536-647-1

Query Match 99.8%; Score 832.4; DB 4; Length 834;
 Best Local Similarity 99.8%; Pred. No. 6,4e-205;
 Matches 833; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 ATGCTTCCGCGGAGGCGGAGACCTGCGAGAGGCGGCGGCGGAGGCGGCGGCGG 60
D 1 ATGCTTCCGCGGAGGCGGAGACCTGCGAGAGGCGGCGGCGGCGGCGGCGGCGG 60
QY 61 CAGCGCCCTTCTGATAGGGGTGAGCGGCGGCACTGCGACGGGAAAGTCGCGTGT 120
D 61 CAGCGCCCTTCTGATAGGGGTGAGCGGCGGCACTGCGACGGGAAAGTCGCGTGT 120
QY 121 GAGAAATCATGAGTGTCTGGGACAGAAAGAGTGAACAGCGGACGGGAAAGTGT 180
D 121 GAGAAATCATGAGTGTCTGGGACAGAAAGAGTGAACAGCGGACGGGAAAGTGT 180
QY 181 ATCTGAGCCAGGACAGGTTCTACAGAGTCTGACGGGACAGAAAGGCGGCTTG 240
D 181 ATCTGAGCCAGGACAGGTTCTACAGAGTCTGACGGGACAGAAAGGCGGCTTG 240
QY 241 AAAGGACAGTACATTTTGAACATCCAGATGCTTGTATATGATTTGATGACAGGACT 300
D 241 AAAGGACAGTACATTTTGAACATCCAGATGCTTGTATATGATTTGATGACAGGACT 300
QY 301 CTGAAGAACATCTGAGAGGCGAAACGATGAGGTGCGACCTATGATTTTGTGACAC 360
D 301 CTGAAGAACATCTGAGAGGCGAAACGATGAGGTGCGACCTATGATTTTGTGACAC 360
QY 361 TCAAGGTTACAGAGACCGGTGTCTACCTTGGGACGTGTCTTGTGAGGCGATC 420
D 361 TCAAGGTTACAGAGACCGGTGTCTACCTTGGGACGTGTCTTGTGAGGCGATC 420
QY 421 TTGGTGTCTACAGGACGAGATCCGGGACATGTTCCAGCTGCGCTCTTGTGAGGCGATC 480
D 421 TTGGTGTCTACAGGACGAGATCCGGGACATGTTCCAGCTGCGCTCTTGTGAGGCGATC 480
QY 481 GACTCCGACGTCAGGCTGTCTGAAAGTTCGCGGACGTGCGCGAGGAGGAGCTG 540
D 481 GACTCCGACGTCAGGCTGTCTGAAAGTTCGCGGACGTGCGCGAGGAGGAGCTG 540
QY 541 GAGAGATTCAGAGGACGATACCACTTGTGAAAGCGGCTTGTGAGGAGTTCGCTG 600
D 541 GAGAGATTCAGAGGACGATACCACTTGTGAAAGCGGCTTGTGAGGAGTTCGCTG 600
QY 601 CCGCAAGAAAGTATGCGGATGTATCATCCACAGAGAGTGAACAATATGTTGCCATC 660
D 601 CCGCAAGAAAGTATGCGGATGTATCATCCACAGAGAGTGAACAATATGTTGCCATC 660
QY 661 AACCTGATGTCAGACATCCAGAGACATTCGATGATGATCATCTGCAAAATGCGAC 720
D 661 AACCTGATGTCAGACATCCAGAGACATTCGATGATGATCATCTGCAAAATGCGAC 720
QY 721 GAGAGGTCATGAGGCGGAGCTTCAAGGAGACCTTTTGTGAGGCGGAGACCTTGG 780
D 721 GAGAGGTCATGAGGCGGAGCTTCAAGGAGACCTTTTGTGAGGCGGAGACCTTGG 780
QY 781 ATGCTGACCTCTGCAAAAGGTCATTTGAGATCCAGAGACAGACCCCACTGA 834
D 781 ATGCTGACCTCTGCAAAAGGTCATTTGAGATCCAGAGACAGACCCCACTGA 834

```

RESULT 3

```

US-09-221-017B-368
; Sequence 368, Application US/09221017B
; Patent No. 6444799
; GENERAL INFORMATION:
; APPLICANT: Ross, Bruce C.
; TITLE OF INVENTION: P. GINGIVALIS NUCLEOTIDES AND USES THEREOF
; NUMBER OF SEQUENCES: 1120
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER

```



```

: STREET: 755 PAGE MILL ROAD
: CITY: Palo Alto
: STATE: CA
: COUNTRY: USA
: ZIP: 94304-1018
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: Windows
:
: SOFTWARE: FastSeq for Windows Version 2.0b
:
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/221,017B
: FILING DATE: 23-DEC-1998
:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: PPI182
: FILING DATE: 31-DEC-1997
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: PPI546
: FILING DATE: 30-JAN-1998
: PRIOR APPLICATION DATA: PPI2911
: FILING DATE: 09-APR-1998
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: PCT/AU98/01023
: FILING DATE: 10-DEC-1998
: ATTORNEY/AGENT INFORMATION:
: NAME: MONTROY, Gladys H
: REGISTRATION NUMBER: 32,430
: REFERENCE/DOCKET NUMBER: 27340-20021.00
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 650-813-5600
: TELEFAX: 650-494-0792
:
: TELEX: 706141
:
: INFORMATION FOR SEQ ID NO: 368:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 5687 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: circular
: MOLECULE TYPE: DNA (genomic)
: HYPOTHEetical: NO
: ANTI-SENSE: UNKNOWN
: ORIGINAL SOURCE:
: ORGANISM: PORPHYROMONAS GINGIVALIS
:
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: 1...5687
:
: US-09-221-017B-368
:
: Query Match 12.8%; Score 106.8; DB 4; Length 5687;
: Best Local Similarity 54.3%; Pred. No. 4.1e-18;
: Matches 238; Conservative 0; Mismatches 197; Indels 3; Gaps 1;
:
: QY 248 AGTACATTTTGAACATCCAGATGCTTGTATATGATGACAGGACTGTGAAGA 307
: DB 4854 AGCTGAACCTACGACATCCCAATGCTTTGACACGAGTATGTCGTAGGATATCTCT 4913
: QY 308 ACATCGTGAAGGCAAAACGCTGAGGTGCGGACCTATGATTTTGTGACACTCAAGT 367
: DB 4914 CTCTGAAGGAGGCAAAACGATAGAGGTGCGGCTATCTTTCTGTAGACACAATCGTT 4973
: QY 368 TACCAAGACCAAGGTGTCTACCTCGGACGCTGTTCTGTTGAGGGCATCTTGCTGT 427
: DB 4974 TGAAGAAAAAGTAAACCGTTCCTGTCACAAAGTATCGTACGTGATGATCTGATAT 5033
: QY 428 TCTACAGCCAGAGATCCGGGACATGTTCCACCTGCGGCTCTTGTGAGACACCGACCG 487
: DB 5034 TCGAAGAACAAAGAGCTCGGGATCTGATGATGTGAAGATTCGTGATACCGATGCG 5093
: QY 488 ACCTCAGGCTGTCTGAAAGATTCCTCGGAGCT---GCGCCGAGGAGGAGCACTTGAGC 544
: DB 5094 ATATTGATTTGGCGCGCGCTTGTGCGGATGTCCAGGAACGGGAGCGCAATATGAT 5153

```

```

: QY 545 AGATCTGACGAGTACACCACTTGTGTAAGCGGCTTTCGAGGAGTTCTGCTGCCGA 604
: DB 5154 CGGTATTGGACATATCTTCAAGCTTTCGCGCTTTCGACGAGGATTTGTGGAACCAT 5213
: QY 605 CAAAGAGTATGCGGATGTGATCATCCACAGAGAGTGAACATATGTTGCGCATCAACC 664
: DB 5214 CCAAGGGTATGCGGATGTGATCATTCGGAAGGTGATTCATTTGGTGGGCTCTCAC 5273
: QY 665 TGATCTGACGACATCC 682
: DB 5274 TCTTGTGCAAAAAATCC 5291
:
: RESULT 4
: US-08-961-527-76/c
: Sequence 76, Application US/08961527
: Patent No. 6420135
:
: GENERAL INFORMATION:
: APPLICANT: Charles Kunsch
: TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
: NUMBER OF SEQUENCES: 391
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Human Genome Sciences, Inc.
: STREET: 9410 Key West Avenue
: CITY: Rockville
: STATE: Maryland
: COUNTRY: USA
: ZIP: 20850
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
: COMPUTER: HP Vectra 486/33
: OPERATING SYSTEM: MSDOS version 6.2
: SOFTWARE: ASCII text
:
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/961,527
: FILING DATE:
: CLASSIFICATION: 424
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Brookes, A. Anders
: REGISTRATION NUMBER: 36,373
: REFERENCE/DOCKET NUMBER: PB340P1
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (301) 309-8504
: TELEFAX: (301) 309-8512
:
: INFORMATION FOR SEQ ID NO: 76:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 10011 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
:
: US-08-961-527-76
:
: Query Match 12.2%; Score 101.8; DB 4; Length 10011;
: Best Local Similarity 53.2%; Pred. No. 9.3e-17;
: Matches 239; Conservative 0; Mismatches 207; Indels 3; Gaps 1;
:
: QY 252 CAATTTTGCACATCCAGATGCTTGTATATGATTTGACACAGACTCTGAAGAACT 311
: DB 10001 CAACTATGACATCTTTTGTGATACAGACTGTGATGACAGACTTAAAGAAAT 9942
: QY 312 CGTGAGGCAAAACGCTGAGGTGCGGACCTATGATTTTGTGACACTCAAGTTACC 371
: DB 9941 GTTGGAGGGCGCTCGGTGACATCCGACCTAAGACTATTAAGAGCAACAGAGTGG 9882
: QY 372 AAGAACCAAGGTGTCTACCTTGGCGAGCTGTTCTGTTGAGGCGATCTTGTTCTTA 431
: DB 9881 CAAGACTTATCTGTACAGAACTCAAGATGTCTTTATCTGTTGAGGATTTTGTGGA 9822
: QY 432 CAGCAGAGATCCGGGACATGTTCAACTGCGCTCTTCTGTGACACCGACTCCGAGCT 491

```

```
Db      9821 GGACACAGCTCTGCGCATTTTGATGATCAATATTTTGTGGATTCGATGACGATGT 9762
Qy      492 CAGGCTGTCTCGAAGAGTTCTCCGGGA---CGTGGCCGAGGAGGAGGAGCTGAGCAGAT 548
Db      9761 GCGGATTATTCGTGATCAAGCGTGATGATGAGAGCGTGGCGGTGACCTTGATGACGT 9702
Qy      549 TCTGACGAGTACACACCTTCGTGAAGCCGCGCTTGAGAGAGTTCTGCGCCGAGAAA 608
Db      9701 TATTAAACAGTACTTAGGTGTGTCAAACAAATGATACACCGATTATCGAGTCACTAA 9642
Qy      609 GAATGATGCGGATGATGATCATCCACGAGAGTGAACAATATGTTGCTCACTGAT 668
Db      9641 GCGTATGCTGATATGCTGATCTCTGAAGGGGTTAGCAATACCGTGGCTATGACTGTT 9582
Qy      669 CGTGACGACATCCAGACATTTGAAATG 697
Db      9581 GACGACCAAGATTCGAAAGATTTGGAAAG 9553

RESULT 5
US-09-134-001C-781
; Sequence 781, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 781
; LENGTH: 651
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-781

Query Match      10.3%; Score 86.2; DB 4; Length 651;
Best Local Similarity 51.5%; Pred. No. 4,1e-13;
Matches 224; Conservative 0; Mismatches 208; Indels 3; Gaps 1;

Qy      253 AATTTGACATCCAGATGCTTTGATATGATTTGATGACAGAGACTCTGAAGATC 312
Db      196 AATTATGACCACTTATGATTTGATATGATTTACTATTCTAATTTTAAAGATTTA 255
Qy      313 GTGAGGGCAAAACGGTGGAGGTGCGACTATGATTTTGTGACACACTCAAGTTACA 372
Db      256 AGAATGGAACCAAGTAGAGTACTAGTACATTAATCTCAACATACACGTAGTAAA 315
Qy      373 GAGACACGGGTGCTACCTCCGGAGAGTGTCTGTTGAGGGAGCTGGTGTCTAC 432
Db      316 GAAACAATTGATTTGATTCAAAAGATGTATATCGTAGAGATATTGCTTTGAA 375
Qy      433 AGCCAGAGATCCGAGACATGTTCCACTGCGCTCTTCTGTGACACCGACTCCGAGCTC 492
Db      376 AACAACACATTACGAGACATGATGATGTGAATTTTATGTCATGTCGATGCTGATTTA 435
Qy      493 AGGCTGTCTGAAGAGTTCTCCGGGA---CGTGGCCGAGGAGGAGACTGAGCAGATT 549
Db      436 CGAATTTTACGTAGGCTTACAAAGATACTAAGAGCGTGTCTGTAACAATGGAATCA 495
Qy      550 CTGACGAGTACACACCTTGTGTAAGCGGCGCTTTCAGAGAGTTCTGCTGCGGCAAG 609
Db      496 ATTATATCAATATCTTAATGTAGTAAGACCTATGATGAGCAATTTATTGAACTTA 555
Qy      610 AAGTATGCGGATGATCATCCACGAGAGTGAACAATATGTTGCGATCAACCTGATC 669
Db      556 AAGCATGCTGACATCATATTCTGTAAGAGGAGCAATAAGTTGCGATGATATTATG 615
```

```
Qy      670 GTGACACATCCAG 684
Db      616 ACTACTAAATCCAG 630

RESULT 6
US-09-107-532A-1370
; Sequence 1370, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCIT
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 1370:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 633 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; MOLECULE TYPE: circular
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (b) LOCATION 1...633
; SEQUENCE DESCRIPTION: SEQ ID NO: 1370:
US-09-107-532A-1370

Query Match      8.8%; Score 73.4; DB 4; Length 633;
Best Local Similarity 50.1%; Pred. No. 7.9e-10;
Matches 210; Conservative 0; Mismatches 206; Indels 3; Gaps 1;

Qy      253 AATTTGACATCCAGATGCTTTGATATGATTTGATGACAGAGCTTGAAGAACATC 312
Db      178 AATTAGATCAACCATTTGCTTTGATATGATCTCTGATTCAGCATGAGCGGACTTA 237
Qy      313 GTGAGGGCAAAACGGTGGAGGTGCGACTATGATTTTGTGACACACTCAAGTTACA 372
Db      238 CTGAACATTAAGCCATTGAAAAACCGGTATGACTATGTAGCGCATACAAAGTCAAG 297
Qy      373 GAGACACGGGTGCTACCTCGGAGAGTGTCTGTTGAGGGGCACTTGTGTTCTAC 432
```

Db 298 GCAACGATCATTCAGAACCAAAAAGATGATCATTTAGAGGAATCTGATTTAGAA 357
Qy 433 AGCCAGAGATCCGGGACGATTTCCACCTCTTCTTGACACCCGACTCCGAGTC 492
Db 358 GACGAACGTCGTGGTGAAGTGAATGAATGAATGAATGAATGAATGAATGAATGAAT 417
Qy 493 AGGCTGCTCGAAGAGTTCTCCGGGACGTG--CGCCGAGGAGGAGCTGAGCAGATT 549
Db 418 CGTATCATTTCTGTATCAACCGGATATGAGAGACGGGACGACATTGACTCCGTG 477
Qy 550 CTGACGAGTACACCACTTCGTGAAGCCGCGCTTCGAGAGTTCTGCTGCCGACAAAG 609
Db 478 ATCGAACATATTTGACAGTTGTTAAACGATGATCATCAATTTATCGAACCAAG 537
Qy 610 AAGTATGCCATGTGATCATCCACGAGAGTGACATATGTTGCTGATCAACCTGAT 668
Db 538 CGCTATGCGGATATCATTTGTTCCAGAGGCGAGAAAACACGTTGCTATGATTTGAT 596

RESULT 7
US-09-198-452A-1/C
Sequence 1, Application US/09198452A
Patent No. 6559294
GENERAL INFORMATION:
APPLICANT: Grifffals, R.
TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments thereof and uses thereof, in particular for the diagnosis, prevention of infection
FILE REFERENCE: 9710-003-999
CURRENT APPLICATION NUMBER: US/09/198,452A
CURRENT FILING DATE: 1998-11-24
NUMBER OF SEQ ID NOS: 6849
SEQ ID NO 1
LENGTH: 1230025
TYPE: DNA
ORGANISM: Chlamydia pneumoniae
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)-(15000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (15001)-(30000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (30001)-(45000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (45001)-(60000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (60001)-(75000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (75001)-(90000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (90001)-(105000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (105001)-(120000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (120001)-(135000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (135001)-(150000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (150001)-(165000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (165001)-(180000)
OTHER INFORMATION: n=a or c or g or t

NAME/KEY: misc feature
LOCATION: (180001)-(195000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (195001)-(210000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (210001)-(225000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (225001)-(240000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (240001)-(255000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (255001)-(270000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (270001)-(285000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (285001)-(300000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (300001)-(315000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (315001)-(330000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (330001)-(345000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (345001)-(360000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (360001)-(375000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (375001)-(390000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (390001)-(405000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (405001)-(420000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (420001)-(435000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (435001)-(450000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (450001)-(465000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (465001)-(480000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (480001)-(495000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (495001)-(510000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (510001)-(525000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (525001)-(540000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (540001)-(555000)
OTHER INFORMATION: n=a or c or g or t

LOCATION: (540001)..(555000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (555001)..(570000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (570001)..(585000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (585001)..(600000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (600001)..(615000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (615001)..(630000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (630001)..(645000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (645001)..(660000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (660001)..(675000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (675001)..(690000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (690001)..(705000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (705001)..(720000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (720001)..(735000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (735001)..(750000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (750001)..(765000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (765001)..(780000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (780001)..(795000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (795001)..(810000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (810001)..(825000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (825001)..(840000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (840001)..(855000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (855001)..(870000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (870001)..(885000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (885001)..(900000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (900001)..(915000)

OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
Query Match 6.7%; Score 55.8; DB 4; Length 1230025;
Best Local Similarity 49.9%; Pred. No. 0.00028;
Matches 197; Conservative 0; Mismatches 192; Indels 6; Gaps 2;
Qy 248 AGTACATTTTTCACATCCGATGCTTGATATATTTGATGACACAGACTCTGAAGA 307
Db 835626 ATTTAATTGGGATCATCCGAGCCCTTGATATGATCTATTAATTTACACATAAAG 835567
Qy 308 ACATCGTGAGGCGCAAAACGCTGAGAGTCCGACCTATGATTTTGACA---CACTCAA 364
Db 835566 GTCTAAATAATATGATGATGTCACAGCCCAATTTTGTATTTTATGTAATGAT 835507
Qy 365 GGTTCACGAGACCAAGGTGCTTACCTCGGACGTGTTCTGTTTGAAGGCATCTTG 424
Db 835506 CTAAACGGAGATGAAACGATCTATCCATCTAAAGTTATTTCTGTTGAAGGTATCTGG 835447
Qy 425 TGTTCACGCCGAGAGATCCGGACATGTTCCACCTGCGCCCTTCTGTGACACCGACT 484
Db 835446 TCTTTGAAATCAAGAACTTAGAGATCTTAGATATAGAGATCTTTGTAGACACCGATG 835387
Qy 485 CCGACGTCAAGCTGTCTCGAAGATTCTCCGGACGTGC---GCCGAGGAGGAGACTTG 541
Db 835386 CTGATGAAGAGATACACGCCGATGTTGAGATGTTCAAGAAACAGAGATAGCGTGG 835327
Qy 542 AGCAGATTCTGACGACGATACCACTTGCTGTAAGCCGCGCTTCGAGAGTTCTGCCTGC 601
Db 835326 ACTGCATCATGCTCTGTTATCTTCTTAGTGTAAGCCTATGCAATGAATTTATAGAGC 835267
Qy 602 CGACAAAGAGTATGCCAGTGTATCATCCACGA 636
Db 835266 CGACTCGAATATGCTGATATCATTTGTACATGA 835232
RESULT 8
US-09-557-884-1
Sequence 1, Application US/09557884
Patent No. 6506581
GENERAL INFORMATION:
APPLICANT: Fleischmann et al.
TITLE OF INVENTION: The Nucleotide sequence of the Haemophilus influenzae Rd Genome, Fragments thereof, and Uses Thereof
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: MD
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: 3 1/2 inch diskette
COMPUTER: Dell Pentium
OPERATING SYSTEM: MS DOS V6.22
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/557,884
FILING DATE: 25-Apr-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/476,102
FILING DATE: JUN-5-1995
ATTORNEY/AGENT INFORMATION:
NAME: Michelle S. Marks
REGISTRATION NUMBER: 41,971
REFERENCE/DOCKET NUMBER: PB186P3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 301-309-8504
TELEFAX: 301-309-8439
INFORMATION FOR SEQ ID NO: 1:


```
STATE: NY
COUNTRY: U.S.A.
ZIP: 10591
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/469,537A
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: USSN 08/406,247
FILING DATE: 17-MAR-1995
APPLICATION NUMBER: USSN 08/144,992
FILING DATE: 28-OCT-1993
APPLICATION NUMBER: USSN 07/736,559
FILING DATE: 26-JUL-1991
ATTORNEY/AGENT INFORMATION:
NAME: Kempler, Ph.D., Gail M
REGISTRATION NUMBER: 32,143
REFERENCE/DOCKET NUMBER: REG 070C
TELECOMMUNICATION INFORMATION:
TELEPHONE: 914-345-7400
TELEFAX: 914-345-7721
TELEX:
INFORMATION FOR SEQ ID NO: 106:
SEQUENCE CHARACTERISTICS:
LENGTH: 4092 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
FEATURE:
NAME/KEY: Coding Sequence
LOCATION: 200...3028
OTHER INFORMATION:
NAME/KEY: Human ROR2
LOCATION: 1...4092
OTHER INFORMATION:
US-08-469-537A-106

Query Match 6.0%; Score 50.4; DB 2; Length 4092;
Best Local Similarity 48.9%; Pred. No. 0.0012;
Matches 135; Conservative 0; Mismatches 141; Indels 0; Gaps 0;

QY 129 CATGAGTTGCTGGACAGAAAGAGTGAACAGCGGACGCGAAGGTGTCATCTGAG 188
DB 1552 CATGAAATGCCCTCATTTAACAGACACAAACAGGCCCAAACTCAAAAGATCACTGTC 1611
QY 189 CCAGACAGGTTCTTCAACAGGCTCTGACGGGACAGACAGAGCCCAAGGCTTGAAGACA 248
DB 1612 TGCCTGAGGTTCTATGAGAGGCTGGAGAGGACCGGTTTGGAAAGTCTAACAAAGGTCA 1671
QY 249 GTACAATTTTGAACCATCCAGATGCTCTTGTATATGATTGTGACACAGACTTGAAGAA 308
DB 1672 CCTGTTGGCCCTGCCCGGGGGAGACAGACCCAGGCTGTGGCATCAAAACGCTGAAGGA 1731
QY 309 CATGTGAGGAGGCAAAACGCTGAGAGTCCGACCTATGATTTTGTACACACTCAAGTT 368
DB 1732 CAAGCGAGGAGGCCCTGCGGAGGAGATTCCGGCATGAGGCTATGCTGGAGACGCGCT 1791
QY 369 ACCAGAGCACAGGAGTCTACCTCGGCGAGCTGTT 404
DB 1792 GCAACACCCCAAGCTGCTGCTGCTGGGCTGTGT 1827

RESULT 11
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
```

```
APPLICANT: DORNER, F.
APPLICANT: SCHIEFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: pTZ9pt-Fls
US-08-232-463-14

Query Match 5.5%; Score 46; DB 1; Length 7218;
Best Local Similarity 4.7%; Pred. No. 0.019;
Matches 19; Conservative 216; Mismatches 171; Indels 0; Gaps 0;

QY 71 TCCTGATAGGGGTGACGGCGGACACTGCACGCGGGAAGTCGACCGTGTGAGAAAGATCA 130
DB 1437 TACRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1378
QY 131 TGAAGTTGCTGGACAGAAAGAGTGAACAGCGGACGCGAAGGTGTCATCTGAGCC 190
DB 1377 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1318
QY 191 AGGACAGGTTCTTCAAGGCTCTGACGGGACAGACAGAGGCAAGGCTTGAAGACACT 250
DB 1317 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1258
QY 251 ACAATTTGACCATCCAGATGCTTGTATATGATTGTGACACAGACTTGAAGACA 310
DB 1257 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1198
QY 311 TCGTGAAGGCAAAACGCTGAGAGTCCGACCTATGATTTTGTACACACTCAAGTTAC 370
DB 1197 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1138
QY 371 CAGAGCACAGGAGTCTACCTCGGAGAGTGTTCTGTTTGAAGGACATCTGTTCT 430
DB 1137 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1078
QY 431 ACAGCCAGAGATCCGGACATGTTCACTGCGCTCTTCTGTGA 476
```

Db 1077 RRRRRRRRRATCGCAAGCTCCCTGACCTGCGACCAAGCTGCGA 1032

RESULT 12
US-09-252-991A-931

; Sequence 931, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 931

; LENGTH: 1437

; TYPE: DNA

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-931

Query Match 5.2%; Score 43; DB 4; Length 1437;
Best Local Similarity 52.5%; Pred. No. 0.067;
Matches 94; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

Qy 431 ACAGCCGAGATCCGGGACATGTTCCACCTGCGCCTTCTTGAGACACCGACTCGGACG 490
Db 218 ACAGCTGGCTGTTCCGACCACTACGACTGCTGCTACCGACTTGGCAGCGCGCAAG 277
Qy 491 TCAGCTGTCTCGAAGATTCTCGGAGCGTGGCCGAGGAGGAGACCTGAGCAGATTTC 550
Db 278 GCTGGCGCAACTCCGGGACCTGGCGGACGCAACTCAAGCCGCAAGGGCATCGAGCTGTGG 337
Qy 551 TGACGAGTACACCACTTGTGTAAGCGGCTTTCGAGGATTCTGCTCCGCAAG 609
Db 338 TGTGTACACGAGCCGCGGCGGCTGTGAAACCGCGAAGCTCAGCCCGGAAAG 396

RESULT 13

US-09-252-991A-1020

; Sequence 1020, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 1020

; LENGTH: 1647

; TYPE: DNA

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-1020

Query Match 5.2%; Score 43; DB 4; Length 1647;
Best Local Similarity 52.5%; Pred. No. 0.07;
Matches 94; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

Qy 431 ACAGCCGAGATCCGGGACATGTTCCACCTGCGCCTTCTTGAGACACCGACTCGGACG 490
Db 720 ACAGCTGGCTGTTCCGACCACTACGACTGCTGCTACCGACTTTCGACACCGCGCAAG 779

Qy 491 TCAGCTGTCTCGAAGATTCTCCGGGACGTGGCGCGGAGGAGGACCTGAGCAGATTTC 550
Db 780 GCTGGCGCAACTCCGGGACCTGCGGACGAACTCAAGGCGCAAGGGCATCGAGCTGTGG 839

Qy 551 TGACGAGTACACCACTTGTGTAAGCGGCTTTCGAGGAGATTCTGCTCCGCAAG 609
Db 840 TGTGTACACGAGCCCGCGGCTGTGTAACCGGAGAACTCAGCCCGGAAAG 898

RESULT 14
US-09-252-991A-1036/c

; Sequence 1036, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 1036

; LENGTH: 2847

; TYPE: DNA

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-1036

Query Match 5.2%; Score 43; DB 4; Length 2847;
Best Local Similarity 52.5%; Pred. No. 0.083;
Matches 94; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

Qy 431 ACAGCCGAGATCCGGGACATGTTCCACCTGCGCCTTCTTGAGACACCGACTCGGACG 490
Db 914 ACAGCTGGCTGTTCCGACCACTACGACTGCTGCTACCGACTTGGCAGCGCGCAAG 855
Qy 491 TCAGCTGTCTCGAAGATTCTCCGGAGCGTGGCCGAGGAGGAGACCTGAGCAGATTTC 550
Db 854 GCTGGCGCAACTCCGGGACCTGGCGGACGCAACTCAAGCGCAAGGGCATCGAGCTGTGG 795

Qy 551 TGACGAGTACACCACTTGTGTAAGCGGCTTTCGAGGATTCTGCTCCGCAAG 609
Db 794 TGTGTACACGAGCCGCGGCGGCTGTGTAACCGGAGAACTCAGCCCGGAAAG 736

RESULT 15

US-09-252-991A-12919

; Sequence 12919, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 12919

; LENGTH: 444

; TYPE: DNA

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-12919

Query Match 4.8%; Score 40.4; DB 4; Length 444;
Best Local Similarity 53.1%; Pred. No. 0.22;
Matches 86; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 25, 2003, 02:04:12 : Search time 58.3596 Seconds
(without alignments)
6307.673 Million cell updates/sec

Title: US-09-896-522-3

Sequence: 1 atgagcttcgagcgagcgagcga.....ccagcagcagcagccactga 834

Scoring table: OLIGO_NIC

Gapop 60.0 , Gapext 60.0

Searched: 569978 seqs, 220691566 residues

Word size : 15

Total number of hits satisfying chosen parameters: 425

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

Issued Patents NA: *
1: /cgn2_6/ptodata/2/ina/5A COMB.seq: *
2: /cgn2_6/ptodata/2/ina/5B COMB.seq: *
3: /cgn2_6/ptodata/2/ina/6A COMB.seq: *
4: /cgn2_6/ptodata/2/ina/6B COMB.seq: *
5: /cgn2_6/ptodata/2/ina/PTUS COMB.seq: *
6: /cgn2_6/ptodata/2/ina/backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	834	100.0	1288	4 US-09-620-312D-546	Sequence 546, App
2	783	93.9	834	4 US-09-536-647-1	Sequence 1, Appl
3	21	2.5	3060	4 US-08-098-141-1	Sequence 1, Appl
4	17	2.0	600	4 US-09-252-991A-4602	Sequence 4602, Ap
5	17	2.0	887	3 US-09-276-531-4	Sequence 4, Appl
6	17	2.0	1329	4 US-09-134-001C-2296	Sequence 2296, Ap
7	17	2.0	2493	4 US-09-801-861-1	Sequence 1, Appl
8	17	2.0	2493	4 US-09-252-991A-4346	Sequence 4346, Ap
9	17	2.0	2816	4 US-09-962-665-4	Sequence 4, Appl
10	17	2.0	6297	3 US-09-045-201A-1	Sequence 1, Appl
11	17	2.0	6297	4 US-09-619-062-1	Sequence 1, Appl
12	17	2.0	36412	4 US-08-311-731A-132	Sequence 132, App
13	17	2.0	36412	4 US-08-311-731A-130	Sequence 130, App
14	17	2.0	51719	4 US-09-918-686-2	Sequence 2, Appl
15	17	2.0	53332	4 US-09-801-861-3	Sequence 3, Appl
16	17	2.0	65042	4 US-09-784-316-3	Sequence 3, Appl
17	17	2.0	68750	4 US-09-335-409-1	Sequence 1, Appl
18	17	2.0	68750	4 US-09-568-102-1	Sequence 1, Appl
19	17	2.0	68750	4 US-09-567-969-1	Sequence 1, Appl
20	17	2.0	68750	4 US-09-568-480-1	Sequence 1, Appl
21	17	2.0	68750	4 US-09-568-486-1	Sequence 1, Appl
22	17	2.0	68750	4 US-09-568-472-1	Sequence 1, Appl
23	17	2.0	68750	4 US-09-567-899-1	Sequence 1, Appl
24	17	2.0	71989	4 US-09-443-501A-2	Sequence 2, Appl
25	17	2.0	92139	4 US-09-918-686-1	Sequence 1, Appl
26	17	2.0	4403765	3 US-09-103-840A-2	Sequence 2, Appl
27	17	2.0	4411529	3 US-09-103-840A-1	Sequence 1, Appl

28	16	1.9	185	2 US-08-107-676-27	Sequence 27, Appl
29	16	1.9	270	4 US-09-313-294A-1081	Sequence 1081, Ap
30	16	1.9	350	4 US-08-213-942-2	Sequence 2, Appl
31	16	1.9	369	4 US-09-265-919-10	Sequence 10, Appl
32	16	1.9	432	4 US-09-252-991A-5495	Sequence 5495, Ap
33	16	1.9	432	4 US-09-252-991A-10246	Sequence 10246, A
34	16	1.9	450	4 US-09-252-991A-15499	Sequence 15499, A
35	16	1.9	484	4 US-09-370-838-283	Sequence 283, App
36	16	1.9	511	4 US-09-404-879A-30	Sequence 30, Appl
37	16	1.9	511	4 US-09-338-933-30	Sequence 30, Appl
38	16	1.9	511	4 US-09-215-681-30	Sequence 30, Appl
39	16	1.9	525	1 US-07-924-028A-2	Sequence 2, Appl
40	16	1.9	633	4 US-09-107-532A-1370	Sequence 1370, Ap
41	16	1.9	666	4 US-09-252-991A-15447	Sequence 15447, A
42	16	1.9	740	4 US-09-340-620A-48	Sequence 48, Appl
43	16	1.9	741	4 US-09-252-991A-3445	Sequence 3445, Ap
44	16	1.9	784	3 US-08-998-416-267	Sequence 267, App
45	16	1.9	853	2 US-08-306-511A-5	Sequence 5, Appl

ALIGNMENTS

RESULT 1
US-09-620-312D-546
Sequence 546, Application US/09620312D
Patent No. 6569662
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chenghua
APPLICANT: Aseudi, Vinod
APPLICANT: Zhang, Jie
APPLICANT: Ren, Feiyang
APPLICANT: Chen, Rui-hong
APPLICANT: Zhao, Qing A.
APPLICANT: Wehrman, Tom
APPLICANT: Xue, Aidong J.
APPLICANT: Wang, Yonghong
APPLICANT: Wang, Jian-Rui
APPLICANT: Zhou, Ping
APPLICANT: Ma, Yundong
APPLICANT: Wang, Dunrui
APPLICANT: Wang, Zhilwei
APPLICANT: John Tillinghast
APPLICANT: Drmanac, Radoye T.
TITLE OF INVENTION: No. 6569662e1 Nucleic Acids and
FILE REFERENCE: 784CIP2B
CURRENT APPLICATION NUMBER: US/09/620,312D
CURRENT FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: 09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/488,725
PRIOR FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 1105
SOFTWARE: pt_FL_genes Version 1.0
SEQ ID NO 546
LENGTH: 1288
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (95)..(928)
US-09-620-312D-546

Query Match 100.0%; Score 834; DB 4; Length 1288;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 834; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGCTTCGAGCGAGGCGAGAGCTGCGAGAGCCCGGAGGCGAGCGTCCGAC 60
DB 95 ATGAGCTTCGAGCGAGGCGAGAGCTGCGAGAGCCCGGAGGCGAGCGTCCGAC 154

```
QY 61 CAGGGCCCTTCTGATAGGGGTGAGCGCGGCACTGCGACGGGGAAGTGCACCGTGTGT 120
DB 155 CAGGGCCCTTCTGATAGGGGTGAGCGCGGCACTGCGACGGGGAAGTGCACCGTGTGT 214
QY 121 GAGAAATCATGAGATTGCTGGGACAGAACGAGGTGGAACAGCGGACGCGAAGGTGTGTC 180
DB 215 GAGAAATCATGAGATTGCTGGGACAGAACGAGGTGGAACAGCGGACGCGAAGGTGTGTC 274
QY 181 ATCTGAGCCAGAGACAGGTTCTACAGAGTCTTGAACGAGACAGAAAGGCCAGGCTTGG 240
DB 275 ATCTGAGCCAGAGACAGGTTCTACAGAGTCTTGAACGAGACAGAAAGGCCAGGCTTGG 334
QY 241 AAAGGACAGTCAATTTTGAACATCCAGATGCTTGTGAATGATTTGATGACAGAGACT 300
DB 335 AAAGGACAGTCAATTTTGAACATCCAGATGCTTGTGAATGATTTGATGACAGAGACT 394
QY 301 CTGAAAGAACATCTGAGAGGGCAAAACGATGAGGTGCGACCTTATGATTTTGTGACACAC 360
DB 395 CTGAAAGAACATCTGAGAGGGCAAAACGATGAGGTGCGACCTTATGATTTTGTGACACAC 454
QY 361 TCAAGGTTACAGAGACCAAGTGTCTACCTTCGCGACGTTGTTCTGTTGAGGGCATC 420
DB 455 TCAAGGTTACAGAGACCAAGTGTCTACCTTCGCGACGTTGTTCTGTTGAGGGCATC 514
QY 421 TTGGGTCTTACAGCCAGGAGATCCGGGACATGTTCCACCTGGGCTTCTGTTGAGACAC 480
DB 515 TTGGGTCTTACAGCCAGGAGATCCGGGACATGTTCCACCTGGGCTTCTGTTGAGACAC 574
QY 481 GACTCCGACGTCAGGCTGTCTCGAAGAGTTCTCCGGGACGTCGCGACGAGGGAGGACCTG 540
DB 575 GACTCCGACGTCAGGCTGTCTCGAAGAGTTCTCCGGGACGTCGCGACGAGGGAGGACCTG 634
QY 541 GAGCAGATTCTGACGACGATACACACCTTCTGTAAGCCGCGCTTCTGAGAGTTTGTGCTG 600
DB 635 GAGCAGATTCTGACGACGATACACACCTTCTGTAAGCCGCGCTTCTGAGAGTTTGTGCTG 694
QY 601 CCGCAAAAGAGTATGCGCATGTGATCATCCACGAGAGTGAACAATATGTTGGCATC 660
DB 695 CCGCAAAAGAGTATGCGCATGTGATCATCCACGAGAGTGAACAATATGTTGGCATC 754
QY 661 AACCTGATCTGACGACATCCAGAGCAATTTCTGAATGTGACATCTGCAATGTGACACGA 720
DB 755 AACCTGATCTGACGACATCCAGAGCAATTTCTGAATGTGACATCTGCAATGTGACACGA 814
QY 721 GAGAGGTTCATATGGGCGGAGCTACAGGCGACCTTTTCTGAGCCAGGGAGACACCTTGGG 780
DB 815 GAGAGGTTCATATGGGCGGAGCTACAGGCGACCTTTTCTGAGCCAGGGAGACACCTTGGG 874
QY 781 ATGCTGACCTCTGGCAAAACGGTCACTTTGAGATCCAGACAGACCCCACTGA 834
DB 875 ATGCTGACCTCTGGCAAAACGGTCACTTTGAGATCCAGACAGACCCCACTGA 928

RESULT 2
US-09-536-647-1
; Sequence 1, Application US/09536647
; Patent No. 6579708
; GENERAL INFORMATION:
; APPLICANT: Ho, Yen Sen
; APPLICANT: Johnson, Randall
; TITLE OF INVENTION: No. 6579708el Human Uridine Kinase
; FILE REFERENCE: G950020
; CURRENT APPLICATION NUMBER: US/09/536,647
; CURRENT FILING DATE: 2000-03-27
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 834
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(780)
```

```
US-09-536-647-1
Query Match 93.9%; Score 783; DB 4; Length 834;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 833; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGAGCTTCGGCGGAGGCGGAGACATGCGAGAGCCCGCGCGGAGGACCGACCTCCGAC 60
DB 1 ATGAGCTTCGGCGGAGGCGGAGACATGCGAGAGCCCGCGCGGAGGACCGACCTCCGAC 60
QY 61 CAGCGGCCCTTCTGATAGGGGTGAGCGCGGCACTGCGACGGGGAAGTGCACCGTGTGT 120
DB 61 CAGCGGCCCTTCTGATAGGGGTGAGCGCGGCACTGCGACGGGGAAGTGCACCGTGTGT 120
QY 121 GAGAAATCATGAGATTGCTGGGACAGAACGAGGTGGAACAGCGGACGCGAAGGTGTGTC 180
DB 121 GAGAAATCATGAGATTGCTGGGACAGAACGAGGTGGAACAGCGGACGCGAAGGTGTGTC 180
QY 181 ATCTGAGCCAGAGACAGGTTCTACAGAGTCTTGAACGAGACAGAAAGGCCAGGCTTGG 240
DB 181 ATCTGAGCCAGAGACAGGTTCTACAGAGTCTTGAACGAGACAGAAAGGCCAGGCTTGG 240
QY 241 AAAGGACAGTCAATTTTGAACATCCAGATGCTTGTGAATGATTTGATGACAGAGACT 300
DB 241 AAAGGACAGTCAATTTTGAACATCCAGATGCTTGTGAATGATTTGATGACAGAGACT 300
QY 301 CTGAAAGAACATCTGAGAGGGCAAAACGATGAGGTGCGACCTTATGATTTTGTGACACAC 360
DB 301 CTGAAAGAACATCTGAGAGGGCAAAACGATGAGGTGCGACCTTATGATTTTGTGACACAC 360
QY 361 TCAAGGTTACAGAGACCAAGTGTCTACCTTCGCGACGTTGTTCTGTTGAGGGCATC 420
DB 361 TCAAGGTTACAGAGACCAAGTGTCTACCTTCGCGACGTTGTTCTGTTGAGGGCATC 420
QY 421 TTGGGTCTTACAGCCAGGAGATCCGGGACATGTTCCACCTGGGCTTCTGTTGAGACAC 480
DB 421 TTGGGTCTTACAGCCAGGAGATCCGGGACATGTTCCACCTGGGCTTCTGTTGAGACAC 480
QY 481 GACTCCGACGTCAGGCTGTCTCGAAGAGTTCTCCGGGACGTCGCGACGAGGGAGGACCTG 540
DB 481 GACTCCGACGTCAGGCTGTCTCGAAGAGTTCTCCGGGACGTCGCGACGAGGGAGGACCTG 540
QY 541 GAGCAGATTCTGACGACGATACACACCTTCTGTAAGCCGCGCTTCTGAGAGTTTGTGCTG 600
DB 541 GAGCAGATTCTGACGACGATACACACCTTCTGTAAGCCGCGCTTCTGAGAGTTTGTGCTG 600
QY 601 CCGCAAAAGAGTATGCGCATGTGATCATCCACGAGAGTGAACAATATGTTGGCATC 660
DB 601 CCGCAAAAGAGTATGCGCATGTGATCATCCACGAGAGTGAACAATATGTTGGCATC 660
QY 661 AACCTGATCTGACGACATCCAGAGCAATTTCTGAATGTGACATCTGCAATGTGACACGA 720
DB 661 AACCTGATCTGACGACATCCAGAGCAATTTCTGAATGTGACATCTGCAATGTGACACGA 720
QY 721 GAGAGGTTCATATGGGCGGAGCTACAGGCGACCTTTTCTGAGCCAGGGAGACACCTTGGG 780
DB 721 GAGAGGTTCATATGGGCGGAGCTACAGGCGACCTTTTCTGAGCCAGGGAGACACCTTGGG 780
QY 781 ATGCTGACCTCTGGCAAAACGGTCACTTTGAGATCCAGACAGACCCCACTGA 834
DB 781 ATGCTGACCTCTGGCAAAACGGTCACTTTGAGATCCAGACAGACCCCACTGA 834

RESULT 3
US-08-098-141-1
; Sequence 1, Application US/08098141
; Patent No. 5441875
; GENERAL INFORMATION:
; APPLICANT: Hediger Ph.D., Matthias A.
; APPLICANT: Urea Transporter Polypeptide
; TITLE OF INVENTION: Urea Transporter Polypeptide
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Choate, Hall & Stewart
```

STREET: Exchange Place, 53 State Street
CITY: Boston
STATE: MA
COUNTRY: U.S.A.
ZIP: 02119
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/098,141
FILING DATE: 19930723
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Kaplan Esq., Warren A.
REGISTRATION NUMBER: 34,199
REFERENCE/DOCKET NUMBER: 092662-001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617:227-5020
TELEFAX: 617:227-7566
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3060 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-098-141-1

Query Match 2.5%; Score 21; DB 1; Length 3060;
Best Local Similarity 100.0%; Pred. No. 0.44;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 179 TCATCCTGAGCCAGACAGGT 199
Db 1344 TCATCCTGAGCCAGACAGGT 1364

RESULT 4
US-09-252-991A-4602/c
Sequence 4602, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196, 136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 4602
LENGTH: 600
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-4602

Query Match 2.0%; Score 17; DB 4; Length 600;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 128 TCATGAGTGTGGGGG 144
Db 251 TCATGAGTGTGGGGG 235

RESULT 5

US-09-276-531-4
Sequence 4, Application US/09276531
Patent No. 6183968
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Lal, Preeti
APPLICANT: Hillman, Jennifer L.
APPLICANT: Yue, Henry
APPLICANT: Reddy, Roopa
APPLICANT: Guegler, Karl J.
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF GENES ENCODING
NUMBER OF SEQUENCES: 134
CORRESPONDENCE ADDRESS:
ADDRESSER: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/276,531
FILING DATE: Herewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/079,677
FILING DATE: March 27, 1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Lynn E. Murry, Ph.D.
REGISTRATION NUMBER: 42,918
REFERENCE/DOCKET NUMBER: PA-0008 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 887 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: LUNGPUT03
CLONE: 1472268
US-09-276-531-4

Query Match 2.0%; Score 17; DB 3; Length 887;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 400 GTGGTTCTGTTGAGGG 416
Db 423 GTGGTTCTGTTGAGGG 439

RESULT 6
US-09-134-001C-2296
Sequence 2296, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucet-Re-Stramm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964

;; PRIOR FILING DATE: 1997-11-08
;; PRIOR APPLICATION NUMBER: US 60/055,779
;; PRIOR FILING DATE: 1997-08-14
;; NUMBER OF SEQ ID NOS: 5674
;; SEQ ID NO 2296
;; LENGTH: 1329
;; TYPE: DNA
;; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-2296

Query Match 2.0%; Score 17; DB 4; Length 1329;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 254 ATTTGACCATCCGAT 270
|||
Db 530 ATTTGACCATCCGAT 546

RESULT 7
US-09-801-861-1
; Sequence 1, Application US/09801861
; Patent No. 6492154
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001098
; CURRENT APPLICATION NUMBER: US/09/801,861
; CURRENT FILING DATE: 2001-03-09
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2203
; TYPE: DNA
; ORGANISM: Human
US-09-801-861-1

Query Match 2.0%; Score 17; DB 4; Length 2203;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 782 TGCTGACCTCTGGCAA 798
|||
Db 25 TGCTGACCTCTGGCAA 41

RESULT 8
US-09-252-991A-4346
; Sequence 4346, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 4346
; LENGTH: 2493
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-4346

Query Match 2.0%; Score 17; DB 4; Length 2493;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 128 TCATGAGTTGCTGGGA 144
|||
Db 2293 TCATGAGTTGCTGGGA 2309

RESULT 9
US-09-962-665-4
; Sequence 4, Application US/09962665
; Patent No. 6537759
; GENERAL INFORMATION:
; APPLICANT: Stanton, Jr., Vincent P.
; TITLE OF INVENTION: FOLYPOLYGLUTAMATE SYNTHETASE GENE SEQUENCE
; TITLE OF INVENTION: VARIANCES HAVING UTILITY IN DETERMINING THE
; TITLE OF INVENTION: TREATMENT OF DISEASE
; FILE REFERENCE: 11926-015004
; CURRENT APPLICATION NUMBER: US/09/962,665
; CURRENT FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 09/658,659
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 09/596,033
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 09/357,743
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 09/357,024
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: 60/093,484
; PRIOR FILING DATE: 1998-07-20
; SOFTWARE: FastSeq for Windows Version 4.0
; NUMBER OF SEQ ID NOS: 16
; SEQ ID NO 4
; LENGTH: 2816
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 175..1067
; OTHER INFORMATION: n = g or a
; NAME/KEY: misc_feature
; LOCATION: 341
; OTHER INFORMATION: n = c or g
; NAME/KEY: misc_feature
; LOCATION: 791..1997, 2618, 2653
; OTHER INFORMATION: n = t or c
; NAME/KEY: misc_feature
; LOCATION: 1337
; OTHER INFORMATION: n = c or a
; NAME/KEY: misc_feature
; LOCATION: 2107
; OTHER INFORMATION: nucleotide in position 2107 is g, or absent
; NAME/KEY: misc_feature
; LOCATION: 2583
; OTHER INFORMATION: n = t or g
US-09-962-665-4

Query Match 2.0%; Score 17; DB 4; Length 2816;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 89 GCCGCACTGCCACGGG 105
|||
Db 1460 GCCGCACTGCCACGGG 1476

RESULT 10
US-09-045-201A-1
; Sequence 1, Application US/09045201A
; Patent No. 6110718
; GENERAL INFORMATION:
; APPLICANT: Shisheva, Assia
; TITLE OF INVENTION: No. 6110718el Mammalian Putative
; TITLE OF INVENTION: Phosphatidylinositol-4-Phosphate-5-Kinase
; NUMBER OF SEQUENCES: 2

;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Harness, Dickey & Pierce, P.L.C.
;; STREET: P.O. Box 828
;; CITY: Bloomfield Hills
;; STATE: Michigan
;; COUNTRY: U.S.A.
;; ZIP: 48303
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; OPERATING SYSTEM: IBM PC compatible
;; SOFTWARE: Patentin Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/045,201A
;; FILING DATE:
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Smith, Dean F
;; REFERENCE/DOCKET NUMBER: 4981-098431
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (248) 641-1600
;; TELEFAX: (248) 641-0270
;; INFORMATION FOR SEQ ID NO: 1:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 6297 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: double
;; TOPOLOGY: linear
;; US-09-045-201A-1

Query Match 2.0%; Score 17; DB 3; Length 6297;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 GCGCGAGCGCGAGACT 25
Db 116 GCGCGAGCGCGAGACT 132

RESULT 11
US-09-619-062-1
; Sequence 1, Application US/09619062
; Patent No. 6406875
; GENERAL INFORMATION:
; APPLICANT: Shisheva, Assia
; TITLE OF INVENTION: No. 6406875e1 Mammalian Putative
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Harness, Dickey & Pierce, P.L.C.
; STREET: P.O. Box 828
; CITY: Bloomfield Hills
; STATE: Michigan
; COUNTRY: U.S.A.
; ZIP: 48303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/619,062
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/045,201
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, Dean F
; REFERENCE/DOCKET NUMBER: 4981-098431
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (248) 641-1600
; TELEFAX: (248) 641-0270

;; INFORMATION FOR SEQ ID NO: 1:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 6297 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: double
;; TOPOLOGY: linear
;; US-09-619-062-1

Query Match 2.0%; Score 17; DB 4; Length 6297;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 GCGCGAGCGCGAGACT 25
Db 116 GCGCGAGCGCGAGACT 132

RESULT 12
US-08-311-731A-132/c
; Sequence 132, Application US/08311731A
; Patent No. 6583266
; GENERAL INFORMATION:
; APPLICANT: SMITH, DOUGLAS
; APPLICANT: MAO, JEN-I
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; TITLE OF INVENTION: RELATING TO MYCOBACTERIUM TUBERCULOSIS AND LAPRAE FOR
; NUMBER OF SEQUENCES: 411
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: WOLF, GREENFIELD & SACKS, P.C.
; STREET: 600 ATLANTIC AVENUE
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: USA
; ZIP: 02210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/311,731A
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: GATES, EDWARD R.
; REGISTRATION NUMBER: 31,616
; REFERENCE/DOCKET NUMBER: C0044/7125
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/720-3500
; TELEFAX: 617/720-2441
; INFORMATION FOR SEQ ID NO: 132:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 36412 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: MYCOBACTERIUM LAPRAE
; US-08-311-731A-132

Query Match 2.0%; Score 17; DB 4; Length 36412;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 88 GCGCGACTGCCAGCGG 104
Db 25517 GCGCGACTGCCAGCGG 25501

```
RESULT 13
US-08-311-731A-130/C
; Sequence 130, Application US/08311731A
; Patent No. 6583266
; GENERAL INFORMATION:
; APPLICANT: SMITH, DOUGLAS
; APPLICANT: MAO, JEN-I
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; TITLE OF INVENTION: RELATING TO MYCOBACTERIUM TUBERCULOSIS AND LAPRAE FOR
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 411
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: WOLF, GREENFIELD & SACKS, P.C.
; STREET: 600 ATLANTIC AVENUE
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: USA
; ZIP: 02210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/311,731A
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: GATES, EDWARD R.
; REGISTRATION NUMBER: 31,616
; REFERENCE/DOCKET NUMBER: C0044/7125
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/720-3500
; TELEFAX: 617/720-2441
; INFORMATION FOR SEQ ID NO: 130:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 36941 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: MYCOBACTERIUM LEPRAE
; US-08-311-731A-130

Query Match          2.0%; Score 17; DB 4; Length 36941;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      219 AGACGAGAGGCCAAG 235
Db      13109 AGACGAGAGGCCAAG 13093

RESULT 14
US-09-918-686-2
; Sequence 2, Application US/09918686
; Patent No. 6475739
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary
; APPLICANT: Piroli, Sean
; APPLICANT: Paepfer, Bryan
; APPLICANT: Staehling-Hampton, Karen
; TITLE OF INVENTION: METHODS FOR IDENTIFYING
; TITLE OF INVENTION: GENOMIC DELETIONS
; FILE REFERENCE: 240083.515
; CURRENT APPLICATION NUMBER: US/09/918,686
; CURRENT FILING DATE: 2001-07-30
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
```

```
LENGTH: 51719
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1246, 2572, 2604
; OTHER INFORMATION: n = A,T,C or G
US-09-918-686-2

Query Match          2.0%; Score 17; DB 4; Length 51719;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      290 TGCACAGACTCTGGAAG 306
Db      35965 TGCACAGACTCTGGAAG 35981

RESULT 15
US-09-801-861-3
; Sequence 3, Application US/09801861
; Patent No. 6492154
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: C1001098
; CURRENT APPLICATION NUMBER: US/09/801,861
; CURRENT FILING DATE: 2001-03-09
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 53332
; TYPE: DNA
; ORGANISM: Human
; US-09-801-861-3

Query Match          2.0%; Score 17; DB 4; Length 53332;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      782 TGCTGACCTCTGGCAAA 798
Db      2025 TGCTGACCTCTGGCAAA 2041
```

Search completed: November 25, 2003, 06:54:23
Job time : 68.3596 secs

score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

OM nucleic - nucleic search, using sw model

(without alignments)
10528.307 Million cell updates/sec

Sequence: 1 atggtctcgcgaggcga.....ccagcagcagaccactga 834

Scoring table: IDENTITY_NUC

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

2: gb_hlg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_sts.*
11: gb_sy.*
12: gb_un.*
13: gb_un.*
14: gb_vv.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_nm.*
20: em_on.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_pl.*
25: em_pl.*
26: em_ro.*
27: em_str.*
28: em_un.*
29: em_vl.*
30: em_hg_hum.*
31: em_hg_inv.*
32: em_hg_other.*
33: em_hg_mus.*
34: em_hg_pln.*
35: em_hg_rtd.*
36: em_hg_mam.*
37: em_hg_vrt.*
38: em_sy.*
39: em_hgo_hum.*
40: em_hgo_mus.*
41: em_hgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	834	100.0	834	6	AX449219	Sequence AX449219
2	834	100.0	1624	6	AX449217	Sequence AX449217
3	834	100.0	2160	6	BD157613	Sequence BD157613
4	834	100.0	2160	9	AK022317	Primer fo AK022317
5	832.4	99.8	834	9	AF254133	Homo sapi AF254133
6	832.4	99.8	1022	9	AF231290	Homo sapi AF231290
7	826.4	94.2	1395	9	AF125106	Homo sapi AF125106
8	713	85.5	2228	9	AK057848	Homo sapi AK057848
9	658	78.9	1959	10	BC025146	Mus muscu BC025146
10	634	76.0	2072	9	BC015547	Mus muscu BC015547
11	623.6	74.8	753	6	BD146824	Homo sapi BD146824
12	619.8	74.3	1810	10	MDSURK1	Primer fo L31783
13	591.2	70.9	952	10	MDSURK1A	Mus musculu L31783
14	590.4	70.8	254993	2	AC078885	Mus muscu AC078885
15	566	67.9	192336	2	AC118474	Mus muscu AC118474
16	566	67.9	211075	10	AC076974	Mus muscu AC076974
17	483.6	58.0	734	6	AX540411	Sequence AX540411
18	374	44.8	744	9	AB062451	Homo sapi AB062451
19	374	44.8	1121	9	AF236637	Homo sapi AF236637
20	374	44.8	1209	9	BC002906	Homo sapi BC002906
21	374	44.8	1322	6	AX135546	Sequence AX135546
22	364.6	43.7	1784	5	BC045968	Danio rer BC045968
23	362.8	43.5	1312	10	BC023789	Mus muscu BC023789
24	361.2	43.3	1354	10	AF236636	Mus muscu AF236636
25	300.4	36.0	1467	3	AY119583	Drosophil AY119583
26	251.8	30.2	841	11	BV017303	Drosophil BV017303
27	183	21.9	147492	9	AL358781	Human DNA AL358781
28	182.2	21.8	831	9	D78335	Human DNA D78335
29	177.2	21.2	1847	6	AX463241	Sequence AX463241
30	177.2	21.2	1847	9	AK000524	Sequence AK000524
31	175.8	21.1	1875	9	BC033078	Homo sapi BC033078
32	175.6	21.1	1813	6	AX405842	Sequence AX405842
33	159.6	19.1	1823	10	BC016535	Mus muscu BC016535
34	155.6	18.7	156905	2	AC020806	Drosophil AC020806
35	155.6	18.7	199016	3	AC008204	Drosophil AC008204
36	155.6	18.7	206741	3	AB003747	Drosophil AB003747
37	153.2	18.4	314	10	AB030700	Rattus no AB030700
38	150.2	18.0	336	9	BT006860	Homo sapi BT006860
39	150.2	18.0	336	12	BT007558	Synthetic BT007558
40	145.2	17.4	3246	3	AY129436	Drosophil AY129436
41	143	17.1	233210	2	AC097693	Rattus no AC097693
42	143	17.1	244105	2	AC098897	Rattus no AC098897
43	140.8	16.9	241882	10	AL898027	Mouse DNA AL898027
44	140.8	16.9	254677	2	AC068494	Mus muscu AC068494
45	133	15.9	655	11	BV077340	Drosophil BV077340

ALIGNMENTS

RESULT 1	AX449219				
LOCUS	AX449219	834 bp	DNA	linear	PAT 03-JUL-2002
DEFINITION	Sequence 3 from Patent WO0202761.				
ACCESSION	AX449219				
VERSION	AX449219.1	GI:21697996			
KEYWORDS					
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
REFERENCE	1				
AUTHORS	Glucksmann, M.A.				
TITLE	57658, a human uridine kinase and uses thereof				
JOURNAL	Patent: WO 0202761-A 3 10-JAN-2002;				

FEATURES Millennium Pharmaceuticals, Inc. (US)
Location/Qualifiers
1. .834
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 194 a 218 c 265 g 157 t
ORIGIN
Query Match 100.0%; Score 834; DB 6; Length 834;
Best Local Similarity 100.0%; Pred. No. 2.1e-172;
Matches 834; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGCTTCGGGGGAGGCGAAGACTCGGAGAGCCCCCGCGGAGGCCGACCGTCCGAC 60
DB 1 ATGGCTTCGGGGGAGGCGAAGACTCGGAGAGCCCCCGCGGAGGCCGACCGTCCGAC 60
QY 61 CAGCGGCCCTTCTCGATAGGGGGTGAAGCGGCGCACTGCGAGCGGGAAGTCGACCGTGTGT 120
DB 61 CAGCGGCCCTTCTCGATAGGGGGTGAAGCGGCGCACTGCGAGCGGGAAGTCGACCGTGTGT 120
QY 121 GAGAAATCATGAGATTGCTGGGACAGAACGAGGTGAACAGCGGCAAGCGGAGGTGTC 180
DB 121 GAGAAATCATGAGATTGCTGGGACAGAACGAGGTGAACAGCGGCAAGCGGAGGTGTC 180
QY 181 ATCTGAGCCAGGACAGGTTCTCAAGAGTCTGACGCGAGAGCAAGAGCCAGGCTTG 240
DB 181 ATCTGAGCCAGGACAGGTTCTCAAGAGTCTGACGCGAGAGCAAGAGCCAGGCTTG 240
QY 241 AAAGGACAGTCAATTTTGAACATCCAGATCCCTTTGATTAATGATTTGATGACAGAGCT 300
DB 241 AAAGGACAGTCAATTTTGAACATCCAGATCCCTTTGATTAATGATTTGATGACAGAGCT 300
QY 301 CTGAAGAACATGCTGAGAGGCGAAACCGGTGAGGTCGCGACCTATGATTTTGTGACACAC 360
DB 301 CTGAAGAACATGCTGAGAGGCGAAACCGGTGAGGTCGCGACCTATGATTTTGTGACACAC 360
QY 361 TCAAGGTTTACAGAGACCAAGGTCGTCACCTTCGAGCGGAGTTCGTTGAGAGGATC 420
DB 361 TCAAGGTTTACAGAGACCAAGGTCGTCACCTTCGAGCGGAGTTCGTTGAGAGGATC 420
QY 421 TTGGTGTCTACAGCCAGAGATCCGGGACATGTTTCCACTGCGCTCTTCGTGAGACAC 480
DB 421 TTGGTGTCTACAGCCAGAGATCCGGGACATGTTTCCACTGCGCTCTTCGTGAGACAC 480
QY 481 GACTCCGACGTCAGGCTGTCTCGAAGGTTTCCGGGACGTCGCGCGAGGGAGGGACCTG 540
DB 481 GACTCCGACGTCAGGCTGTCTCGAAGGTTTCCGGGACGTCGCGCGAGGGAGGGACCTG 540
QY 541 GAGGAGATTTCAGGACAGTACACACCTTCGTAAGCGGCGCTTCGAGGAGTTCTGCTG 600
DB 541 GAGGAGATTTCAGGACAGTACACACCTTCGTAAGCGGCGCTTCGAGGAGTTCTGCTG 600
QY 601 CCGCAAAAGAGTATGCGATGTATCATCCACGAGAGTGAACAATATGTTGCCATC 660
DB 601 CCGCAAAAGAGTATGCGATGTATCATCCACGAGAGTGAACAATATGTTGCCATC 660
QY 661 AACCTGATCGTGACGACATCTGAGATGTTGATGATGATGATGATGATGATGATGATG 720
DB 661 AACCTGATCGTGACGACATCTGAGATGTTGATGATGATGATGATGATGATGATGATG 720
QY 721 GGAAGGTCCTGAGGAGGAGTCAAGAGGAGCTTTTCTGAGGAGGAGGAGGAGGAGGAGG 780
DB 721 GGAAGGTCCTGAGGAGGAGTCAAGAGGAGCTTTTCTGAGGAGGAGGAGGAGGAGGAGG 780
QY 781 ATGCTGACCTCTGCGAAACGCTCAATTTGAGATGTCAGAGCAGACCCCACTGA 834
DB 781 ATGCTGACCTCTGCGAAACGCTCAATTTGAGATGTCAGAGCAGACCCCACTGA 834

RESULT 2
AX449217
LOCUS AX449217 1624 bp DNA linear PAT 03-JUL-2002

DEFINITION Sequence 1 from Patent WO0202761.
ACCESSION AX449217
VERSION AX449217.1 GI:21697994
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1
AUTHORS Glucksmann, M.A.
TITLE 57658, a human uridine kinase and uses thereof
JOURNAL Patent: WO 0202761-A 1 10-JAN-2002;
Millennium Pharmaceuticals, Inc. (US)
FEATURES
source
1. .1624
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
94. .927
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD37771.1"
/db_xref="GI:21697995"
/translation="MASAGEDCEBPAPDRPHORFLIGVSGTASGKSTVEKIM
ELIGONEBOROKVILSODREFKVLTAEOAKALKGOYNDHPDPAFMDHRTLK
NIVEGTVFVPTVDPVTHRLPETHVYVADVLVPSGLIVFSQIRDMFHLRFPDT
DSVRLSRVLDVRRGRDLEQILTYTTFVYKPAPEEFLPTKRYADVILPBGVDMV
ALNLVQHLQDILINDICKMRHSGNSGRYKRTFSPGHPFLTSGRSHLESSRP
H"
BASE COUNT 354 a 427 c 506 g 337 t
ORIGIN
Query Match 100.0%; Score 834; DB 6; Length 1624;
Best Local Similarity 100.0%; Pred. No. 2e-172;
Matches 834; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGCTTCGGGGGAGGCGAAGACTCGGAGAGCCCCCGCGGAGGCCGACCGTCCGAC 60
DB 94 ATGGCTTCGGGGGAGGCGAAGACTCGGAGAGCCCCCGCGGAGGCCGACCGTCCGAC 153
QY 61 CAGCGGCCCTTCTCGATAGGGGGTGAAGCGGCGCACTGCGAGCGGGAAGTCGACCGTGTGT 120
DB 61 CAGCGGCCCTTCTCGATAGGGGGTGAAGCGGCGCACTGCGAGCGGGAAGTCGACCGTGTGT 120
QY 121 GAGAAATCATGAGATTGCTGGGACAGAACGAGGTGAACAGCGGCAAGCGGAGGTGTC 180
DB 121 GAGAAATCATGAGATTGCTGGGACAGAACGAGGTGAACAGCGGCAAGCGGAGGTGTC 180
QY 181 ATCTGAGCCAGGACAGGTTCTCAAGAGTCTGACGCGAGAGCAAGAGCCAGGCTTG 240
DB 181 ATCTGAGCCAGGACAGGTTCTCAAGAGTCTGACGCGAGAGCAAGAGCCAGGCTTG 240
QY 241 AAAGGACAGTCAATTTTGAACATCCAGATCCCTTTGATTAATGATTTGATGACAGAGCT 300
DB 241 AAAGGACAGTCAATTTTGAACATCCAGATCCCTTTGATTAATGATTTGATGACAGAGCT 300
QY 301 CTGAAGAACATGCTGAGAGGCGAAACCGGTGAGGTCGCGACCTATGATTTTGTGACACAC 360
DB 301 CTGAAGAACATGCTGAGAGGCGAAACCGGTGAGGTCGCGACCTATGATTTTGTGACACAC 360
QY 361 TCAAGGTTTACAGAGACCAAGGTCGTCACCTTCGAGCGGAGTTCGTTGAGAGGATC 420
DB 361 TCAAGGTTTACAGAGACCAAGGTCGTCACCTTCGAGCGGAGTTCGTTGAGAGGATC 420
QY 421 TTGGTGTCTACAGCCAGAGATCCGGGACATGTTTCCACTGCGCTCTTCGTGAGACAC 480
DB 421 TTGGTGTCTACAGCCAGAGATCCGGGACATGTTTCCACTGCGCTCTTCGTGAGACAC 480
QY 481 GACTCCGACGTCAGGCTGTCTCGAAGGTTTCCGGGACGTCGCGCGAGGGAGGGACCTG 540
DB 481 GACTCCGACGTCAGGCTGTCTCGAAGGTTTCCGGGACGTCGCGCGAGGGAGGGACCTG 540
QY 541 GAGGAGATTTCAGGACAGTACACACCTTCGTAAGCGGCGCTTCGAGGAGTTCTGCTG 600
DB 541 GAGGAGATTTCAGGACAGTACACACCTTCGTAAGCGGCGCTTCGAGGAGTTCTGCTG 600

|||||
Db 634 GAGAGATTCTGACGCGATACACACCTTGGTGAAGCGGCTTCGAGAGATTCTGCTG 693
|||
Qy 601 CCGACAAAGAAGATATGCCGATGTGATCATCCACGAGAGTGCACAAATATGTTGCCATC 660
|||
Db 694 CCGACAAAGAAGATATGCCGATGTGATCATCCACGAGAGTGCACAAATATGTTGCCATC 753
|||
Qy 661 AACCTGATCGGCGACATCCAGAGACATCTGAAATGTGACATCTGCAATGTGACACCGA 720
|||
Db 754 AACCTGATCGGCGACATCCAGAGACATCTGAAATGTGACATCTGCAATGTGACACCGA 813
|||
Qy 721 GAGGGGTCCATGGGCGGAGCTACAGCGGACCTTTCTGAGCGGAGGACCACTCTGGG 780
|||
Db 814 GAGGGGTCCATGGGCGGAGCTACAGCGGACCTTTCTGAGCGGAGGACCACTCTGGG 873
|||
Qy 781 ATGCTGACCTCTGCGAAACGGTCACTTTGGAGTCCGACGACGACCCCACTGA 834
|||
Db 874 ATGCTGACCTCTGCGAAACGGTCACTTTGGAGTCCGACGACGACCCCACTGA 927
|||
RESULT 3
LOCUS BD157613 2160 bp DNA linear PAT 17-JAN-2003
DEFINITION Primer for synthesizing full-length cDNA and use thereof.
ACCESSION BD157613
VERSION BD157613.1 GI:27863371
KEYWORDS JP 2002191363-A/12456.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 2160)
Oca,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
Primer for synthesizing full-length cDNA and use thereof
Patent: JP 2002191363-A 12456 09-JUL-2002;
HELIX RESEARCH INSTITUTE
OS Homo sapiens (human)
PN JP 2002191363-A/12456
PD 09-JUL-2002 JP 2000280990
PF 28-JUL-2000 JP 2000280990
PI TOSHIO OTA,TAKAO ISOGAI,TERUSO NISHIKAWA,KOJI HAYASHI,KOORU
PI SAITO,
PI JUNICHI YAMAMOTO,SHIZUKO ISHII,TOMOYASU SUGIYAMA,AI WAKAMATSU,
PI KEIICHI NAGAI,TERUSUI OTSUKI
PC
C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/ PC
10', C12P21/02,C1201/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof FH Key
FT CDS Location/Qualifiers
1.2160
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 457 a 591 c 671 g 441 t
ORIGIN
Query Match 100.0%; Score 834; DB 6; Length 2160;
Best Local Similarity 100.0%; Pred. No. 2e-112;
Matches 834; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

|||||
Db 215 GAGAAATCATGAGATTGCTGGGACAGAACGAGTGGAAACGGGCGACCGGAAGTGTGTC 274
|||
Qy 181 ATCTTAGCCAGGACAGGTTCTTACAAAGTCTTGAACGCGACAGACGAAAGGCCCTTG 240
|||
Db 275 ATCTTAGCCAGGACAGGTTCTTACAAAGTCTTGAACGCGACAGACGAAAGGCCCTTG 334
|||
Qy 241 AAGGACAGTACAAATTTTGAACCATCCAGATGCTTTGATTAATGATTTGATGACAGACT 300
|||
Db 335 AAGGACAGTACAAATTTTGAACCATCCAGATGCTTTGATTAATGATTTGATGACAGACT 394
|||
Qy 301 CTGAAGAACATCGTGAAGGGCAAAAGCGTGAAGTCCGACCTATGATTTTGAACAC 360
|||
Db 395 CTGAAGAACATCGTGAAGGGCAAAAGCGTGAAGTCCGACCTATGATTTTGAACAC 454
|||
Qy 361 TCAGAGTTACAGAGACCAAGGTGTCTTACCTCGGACGTTGTTTGAAGGCATC 420
|||
Db 455 TCAGAGTTACAGAGACCAAGGTGTCTTACCTCGGACGTTGTTTGAAGGCATC 514
|||
Qy 421 TTGGTGTCTTACAGCCAGAGATCCGGGACATGTTCCACTGGCCTTCTCGTGAAC 480
|||
Db 515 TTGGTGTCTTACAGCCAGAGATCCGGGACATGTTCCACTGGCCTTCTCGTGAAC 574
|||
Qy 481 GACTCCGAGCTGAGGCTGTCTCGAAGATTTCTCCGGAAGTGGCGGAGGAGGACCTG 540
|||
Db 575 GACTCCGAGCTGAGGCTGTCTCGAAGATTTCTCCGGAAGTGGCGGAGGAGGACCTG 634
|||
Qy 541 GAGCAGATTCTGACGCGATACACCACTTGTGAAGCCGGCCTTCGAGGAGTTCTGCTG 600
|||
Db 635 GAGCAGATTCTGACGCGATACACCACTTGTGAAGCCGGCCTTCGAGGAGTTCTGCTG 694
|||
Qy 601 CCGACAAAGAAGATATGCCGATGTGATCATCCACGAGAGTGCACAAATATGTTGCCATC 660
|||
Db 695 CCGACAAAGAAGATATGCCGATGTGATCATCCACGAGAGTGCACAAATATGTTGCCATC 754
|||
Qy 661 AACCTGATCGTGAAGACATCCAGACATTTGAAATGTGACATCTGCAATGTGACACCGA 720
|||
Db 755 AACCTGATCGTGAAGACATCCAGACATTTGAAATGTGACATCTGCAATGTGACACCGA 814
|||
Qy 721 GAGGGGTCCATGGGCGGAGCTTACAAAGCGGACCTTTCTGAGCGGAGGACCACTCTGG 780
|||
Db 815 GAGGGGTCCATGGGCGGAGCTTACAAAGCGGACCTTTCTGAGCGGAGGAGGACCACTCTGG 874
|||
Qy 781 ATGCTGACCTCTGCGAAACGGTCACTTTGGAGTCCGACGACGACCCCACTGA 834
|||
Db 875 ATGCTGACCTCTGCGAAACGGTCACTTTGGAGTCCGACGACGACCCCACTGA 928
|||
RESULT 4
LOCUS AK022317 2160 bp mRNA linear PRI 01-AUG-2002
DEFINITION Homo sapiens CDNA FLJ12255 f1s, clone MAMMA1001476, highly similar
to URIDINE KINASE (EC 2.7.1.48).
ACCESSION AK022317
VERSION AK022317.1 GI:10433687
KEYWORDS oligo capping, f1s (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1
Isogai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y.,
Nishikawa,T., Nagai,K., Sugano,S., Shiratori,A., Sudo,H.,
Wagatsuma,M., Hosokiri,T., Kaku,Y., Kodaira,H., Kondo,H.,
Sugawara,M., Takahashi,M., Chiba,Y., Ishida,S., Murakawa,K.,
Ono,Y., Takiguchi,S., Matanabe,S., Kilmura,K., Murakami,K.,
Ishii,S., Kawai,Y., Saito,K., Yamamoto,J., Wakamatsu,A.,
Nakamura,Y., Nagahara,K., Masuho,Y., Ninomiya,K. and Iwayanagi,T.
NEDO human cDNA sequencing project
Unpublished
JOURNAL
REFERENCE 2 (bases 1 to 2160)
AUTHORS Isogai,T. and Otsuki,T.
TITLE Direct Submission

JOURNAL Submitted (23-AUG-2000) Takao Isogai, Helix Research Institute, Genomics Laboratory, 1532-3 Yana, Kiseazuru, Chiba 297-0812, Japan (E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)

COMMENT NEBO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology; cDNA library construction, 5' - & 3' - end one pass sequencing and clone selection: Helix Research Institute (supported by Japan Key Technology Center etc.) and Department of Virology, Institute of Medical Science, University of Tokyo.

FEATURES
 source
 1. 2160
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MAMMA101476"
 /issue_type="Mammary gland"
 /clone_lib="MAMMA1"
 /note="Cloning vector: pME18FL3"
 95 .928
 /note="unamed protein product"
 /codon_start=1
 /protein_id="BAB14010.1"
 /db_xref="GI:10433688"
 /translation="MASAGBDESPAPADRPHPRLIGVSGTASGSKTYCEKIM ELIGONEVORORRVILSDRFYKVLTAQKAKALGQVNDHPAFNDLIMHRLTK NIVEGKTEVEPTYPYVTHSRLEPTVVPADVLPFGILVPSOEIRDMHRLPVD DSDVRLSRVLRVDRGRDLBOILQYTFVFPAPFEELPKTKYADVILIPGVDMV AINLIYQHIQDILNBDICMHRGSGNSRKYKTFSPGDHMLTSGKSHLESSRP H"

CDS
 457 a 591 c 671 g 441 t

BASE COUNT 457 a 591 c 671 g 441 t

ORIGIN

Query Match 100.0%; Score 834; DB 9; Length 2160;
 Best Local Similarity 100.0%; Pred. No. 2e-172;
 Matches 834; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGGCTTCGGCGGAGCGAAGACTGCGAGAGCCCGCGCGGACCGTCCGAC 60
 95 ATGGCTTCGGCGGAGCGAAGACTGCGAGAGCCCGCGCGGACCGTCCGAC 154
 61 CAGCGGCCCTTCTGATAGGGGTGAGCGCGGCACTGCCAGGGAAAGTCAGCTGTGT 120
 155 CAGCGGCCCTTCTGATAGGGGTGAGCGCGGCACTGCCAGGGAAAGTCAGCTGTGT 214
 121 GAGAAATCATGAGATTCTGGGACAGAACAGAGTGGAAACGGCGAGCGGAAGTGTCTC 180
 215 GAGAAATCATGAGATTCTGGGACAGAACAGAGTGGAAACGGCGAGCGGAAGTGTCTC 274
 181 ATCCGAGCCGAGCAGGTTCTACAAAGTCTCTGACGGAGAGCAGAAAGCCCAAGCTTG 240
 275 ATCCGAGCCGAGCAGGTTCTACAAAGTCTCTGACGGAGAGCAGAAAGCCCAAGCTTG 334
 241 AAAGAGCAGTCAATTTTTCATCCAGATCCTTTGATTAATGATTGATGACAGAGACT 300
 335 AAAGAGCAGTCAATTTTTCATCCAGATCCTTTGATTAATGATTGATGACAGAGACT 394
 301 CTGAAGAACATCTGAGAGGGGCAAAAAGGTGAGAGTGCAGACTTATGATTGTGACACAC 360
 395 CTGAAGAACATCTGAGAGGGGCAAAAAGGTGAGAGTGCAGACTTATGATTGTGACACAC 454
 361 TCAAGGTTACAGAGACCAAGGTTCTACCTCGAGAGCGGTTCTGTTGAGGGATC 420
 455 TCAAGGTTACAGAGACCAAGGTTCTACCTCGAGAGCGGTTCTGTTGAGGGATC 514
 421 TTGGTGTTCATCAGCGAGAGATCCGGGACATGTTTCCACTTGCCTCTTCTGTGACACC 480
 515 TTGGTGTTCATCAGCGAGAGATCCGGGACATGTTTCCACTTGCCTCTTCTGTGACACC 574
 481 GACTCCGACGTCAGGCTGTCTGAAAGATTCTCCGGGACGTGCGCCGAGGGAGGACCTG 540
 575 GACTCCGACGTCAGGCTGTCTGAAAGATTCTCCGGGACGTGCGCCGAGGGAGGACCTG 634

QY 541 GAGCAGATTCTGACGAGTACACCACTTCTGTGAGCCGAGCTTCGAGAGATTCTGCTTG 600
 DB 635 GAGCAGATTCTGACGAGTACACCACTTCTGTGAGCCGAGCTTCGAGAGATTCTGCTTG 634
 QY 601 CCGACAAAGAGTATGCCGATGATCATCCACGAGAGTGGACAATATGTTGCCATC 660
 DB 695 CCGACAAAGAGTATGCCGATGATCATCCACGAGAGTGGACAATATGTTGCCATC 754
 QY 661 AACCTGATGTCGAGACATCCAGGACATTCGAAATGAGACATTCGAAATGAGACCA 720
 DB 755 AACCTGATGTCGAGACATCCAGGACATTCGAAATGAGACATTCGAAATGAGACCA 814
 QY 721 GAGAGGTCATGAGCGAGCTCAACGAGACCTTTCTGAGCCAGGAGACCACTTGAG 780
 DB 815 GAGAGGTCATGAGCGAGCTCAACGAGACCTTTCTGAGCCAGGAGACCACTTGAG 874
 QY 781 ATGCTGACCTTGGCAACCGTCACTTTGAGACTCCAGACGACACCCCACTGA 834
 DB 875 ATGCTGACCTTGGCAACCGTCACTTTGAGACTCCAGACGACACCCCACTGA 928

RESULT 5
 AF254133
 LOCUS
 DEFINITION Homo sapiens uridine kinase mRNA, complete cds.
 ACCESSION AF254133
 VERSION AF254133.1 GI:13924749
 KEYWORDS
 SOURCE
 ORGANISM Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 834)
 Ho, Y.S. and Johnson, R.K.
 Human uridine kinase from prostate cancer cell line (LNCap)
 2 (bases 1 to 834)
 Ho, Y.S. and Johnson, R.K.
 Direct Submision
 Submitted (10-APR-2000) Oncology Research, SmithKline Beecham, 709 Swedeland Road, King of Prussia, PA 19406, USA
 Location/Qualifiers
 1. 834
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 1. 834
 /note="protein kinase"
 /codon_start=1
 /product="uridine kinase"
 /protein_id="AAK49122.1"
 /db_xref="GI:13924750"

CDS
 1. 834
 /note="protein kinase"
 /codon_start=1
 /product="uridine kinase"
 /protein_id="AAK49122.1"
 /db_xref="GI:13924750"

FEATURES
 source
 1. 834
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 1. 834
 /note="protein kinase"
 /codon_start=1
 /product="uridine kinase"
 /protein_id="AAK49122.1"
 /db_xref="GI:13924750"

BASE COUNT 194 a 219 c 265 g 156 t

ORIGIN

Query Match 99.8%; Score 832.4; DB 9; Length 834;
 Best Local Similarity 99.8%; Pred. No. 4.7e-172;
 Matches 833; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 ATGGCTTCGGCGGAGCGAAGACTGCGAGAGCCCGCGCGGACCGTCCGAC 60
 DB 1 ATGGCTTCGGCGGAGCGAAGACTGCGAGAGCCCGCGCGGACCGTCCGAC 60
 QY 61 CAGCGGCCCTTCTGATAGGGGTGAGCGCGGCACTGCCAGGGAAAGTCAGCTGTGT 120
 DB 61 CAGCGGCCCTTCTGATAGGGGTGAGCGCGGCACTGCCAGGGAAAGTCAGCTGTGT 120

```

Qy 121 GAGAAATCATGAGTGTCTGGAGCAGAACAGAGTGGAAACAGCGGACGGAAGTGTCTC 180
Db 121 GAGAAATCATGAGTGTCTGGAGCAGAACAGAGTGGAAACAGCGGACGGAAGTGTCTC 180
Qy 181 ATCTGAGCAGAGCAGAGTCTCAACAGTCTCTGACGAGAGCAGAAAGGCCAAGCCTTG 240
Db 181 ATCTGAGCAGAGCAGAGTCTCAACAGTCTCTGACGAGAGCAGAAAGGCCAAGCCTTG 240
Qy 241 AAAGGACAGTACATTTTGTACCATCCAGTCCCTTTGATATGATTTGATGACAGACT 300
Db 241 AAAGGACAGTACATTTTGTACCATCCAGTCCCTTTGATATGATTTGATGACAGACT 300
Qy 301 CTGAAGAACATCGTGGAGGGCAAAAGAGTGGAGTGCAGCTATGATTTTGTGACACAC 360
Db 301 CTGAAGAACATCGTGGAGGGCAAAAGAGTGGAGTGCAGCTATGATTTTGTGACACAC 360
Qy 361 TCAAGGTTACAGAGACCAACGAGTGTCTACCTGCGGACGAGTGTCTGTTTGAAGGATC 420
Db 361 TCAAGGTTACAGAGACCAACGAGTGTCTACCTGCGGACGAGTGTCTGTTTGAAGGATC 420
Qy 421 TTGGTGTCTTACACGCGAGAGATCCGGACATGTTCCACTGCGCTCTCTGTTGACACC 480
Db 421 TTGGTGTCTTACACGCGAGAGATCCGGACATGTTCCACTGCGCTCTCTGTTGACACC 480
Qy 481 GACTCCGACGTCAGGCTGTCTCGAAGAGTCTCCGGACGTCGCGCGAGGGAGGACCTG 540
Db 481 GACTCCGACGTCAGGCTGTCTCGAAGAGTCTCCGGACGTCGCGCGAGGGAGGACCTG 540
Qy 541 GAGCAGATTCGACGACGATACACCACTTCGTGAAAGCGGCTTCGAGAGTTCGCTG 600
Db 541 GAGCAGATTCGACGACGATACACCACTTCGTGAAAGCGGCTTCGAGAGTTCGCTG 600
Qy 601 CCGACAAAGAAATATGCCGATGTGATATCCACAGAGAGTGGACAAATATGTTGCCATC 660
Db 601 CCGACAAAGAAATATGCCGATGTGATATCCACAGAGAGTGGACAAATATGTTGCCATC 660
Qy 661 AACCTGATCGAGCAGCATCCAGAGACATTCGAAATGTGACATCTGCAATATGGACCGA 720
Db 661 AACCTGATCGAGCAGCATCCAGAGACATTCGAAATGTGACATCTGCAATATGGACCGA 720
Qy 721 GAGAGGTCATAGGCGGAGCTACACAGCAGACCTTTCTGAGCGAGGGACCACTCTGG 780
Db 721 GAGAGGTCATAGGCGGAGCTACACAGCAGACCTTTCTGAGCGAGGGACCACTCTGG 780
Qy 781 ATGCTGACCTCTGGCAAAAGCGTCAATTTGAGATCCAGACAGACCCCACTGA 834
Db 781 ATGCTGACCTCTGGCAAAAGCGTCAATTTGAGATCCAGACAGACCCCACTGA 834

RESULT 6
AF237290 1022 bp mRNA linear PRI 02-AUG-2001
LOCUS Homo sapiens uridine-cytidine kinase 1 (UCK1) mRNA, complete cds.
DEFINITION AF237290
ACCESSION AF237290.1 GI:13506764
VERSION AF237290.1
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1022)
AUTHORS Van Rompay,A.R., Norden,A., Linden,K., Johansson,M. and Karlsson,A.
TITLE Phosphorylation of uridine and cytidine nucleoside analogs by two
human uridine-cytidine kinases
Mol. Pharmacol. 59 (5), 1181-1186 (2001)
JOURNAL MEDLINE
MEDLINE 21203813
PUBMED 11306702
REFERENCE 2 (bases 1 to 1022)
AUTHORS Van Rompay,A.R., Linden,K., Norden,A., Zhu,C., Zheng,X.,
Johansson,M. and Karlsson,A.
TITLE Human uridine-cytidine kinase 1 and 2 : rate limiting enzymes
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 1022)

```

```

AUTHORS Van Rompay,A.R., Linden,K., Norden,A., Zhu,C., Zheng,X.,
Johansson,M. and Karlsson,A.
TITLE Direct Submission
JOURNAL Submitted (19-FEB-2000) IMPI, Clinical Virology, Huddinge
University Hospital, Stockholm 14186, Sweden
FEATURES
source
1..1022
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1..1022
/gene="UCK1"
5..838
/gene="UCK1"
/function="phosphorylates uridine and cytidine to uridine
monophosphate and cytidine monophosphate"
/product="uridine-cytidine kinase 1"
/protein_id="AAK28324.1"
/db_xref="GI:13506765"
/translation="MASAGDEDESPAPADPHORPFLIGVSGTSKSTYCEKIM
ELLGQNEVEORQRKVILSQDRFYKVLAEQAKALKQYDFNDHPAIFNDLHRTLK
NIVEGTVFVPTVDFVTHSRLEPVTVPADAVLPEGLIVFYSOEIRDMFHLFVDT
DSVRLSRVLRDVRGRDLEOILTYTFVPAFEEFLPTKYADVIIPIGVDMV
ATNLIVQHTQDILNGLDICKMHRGSGNRSYKTFSEPDHPMLTSGKSHLESSRP
H"
BASE COUNT 227 a 282 c 324 g 189 t
ORIGIN
Query Match 99.8%; Score 832.4; DB 9; Length 1022;
Best local Similarity 99.9%; Pred. No. 4,7e-172;
Matches 833; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 ATGGCTTCGCGGAGGCGAAGACTGCGAAGCCCCCGCGGAGCGGACGTCGCGAC 60
Db 5 ATGGCTTCGCGGAGGCGAAGACTGCGAAGCCCCCGCGGAGCGGACGTCGCGAC 64
Qy 61 CAGCGGCCCTTCTCTGATTAAGGGGTGAGCGGCGGACCTCCAGCGGAAATCGACCTGTCT 120
Db 65 CAGCGGCCCTTCTCTGATTAAGGGGTGAGCGGCGGACCTCCAGCGGAAATCGACCTGTCT 124
Qy 121 GAGAAATCATGAGTGTCTGGAGCAGAAACAGAGTGGAAACAGCGGACGGAAGTGTCTC 180
Db 125 GAGAAATCATGAGTGTCTGGAGCAGAAACAGAGTGGAAACAGCGGACGGAAGTGTCTC 184
Qy 181 ATCTGAGCAGAGCAGAGTCTCAACAGTCTCTGACGAGAGCAGAAAGGCCCTTG 240
Db 185 ATCTGAGCAGAGCAGAGTCTCAACAGTCTCTGACGAGAGCAGAAAGGCCCTTG 244
Qy 241 AAAGGACAGTACATTTTGTACCATCCAGATGCCCTTTGATATGATTTGATGACAGACT 300
Db 245 AAAGGACAGTACATTTTGTACCATCCAGATGCCCTTTGATATGATTTGATGACAGACT 304
Qy 301 CTGAAGAACATCGTGGAGGGCAAAAGAGTGGAGTGCAGCTATGATTTTGTGACACAC 360
Db 305 CTGAAGAACATCGTGGAGGGCAAAAGAGTGGAGTGCAGCTATGATTTTGTGACACAC 364
Qy 361 TCAAGGTTACAGAGACCAACGAGTGTCTACCTGCGAAGCTGTCTGTTTGAAGGATC 420
Db 365 TCAAGGTTACAGAGACCAACGAGTGTCTACCTGCGAAGCTGTCTGTTTGAAGGATC 424
Qy 421 TTGGTGTCTTACACGCGAGAGATCCGGACATGTTCCACTGCGCTCTTCTGTTGACACC 480
Db 425 TTGGTGTCTTACACGCGAGAGATCCGGACATGTTCCACTGCGCTCTTCTGTTGACACC 484
Qy 481 GACTCCGACGTCAGGCTGTCTCGAAGAGTCTCCGGACGTCGCGCGAGGGAGGACCTG 540
Db 485 GACTCCGACGTCAGGCTGTCTCGAAGAGTCTCCGGACGTCGCGCGAGGGAGGACCTG 544
Qy 541 GAGCAGATTCGACGACGATACACCACTTCGTGAAAGCGGCTTCGAGAGATTCGCTG 600
Db 545 GAGCAGATTCGACGACGATACACCACTTCGTGAAAGCGGCTTCGAGAGATTCGCTG 604

```

QY 601 CCGACAAAGAGTATGCCGATGTGATCATCCACAGAGAGTGAACAATATGTTGCCATC 660
 Db 605 CCGACAAAGAGTATGCCGATGTGATCATCCGAGAGAGTGAACAATATGTTGCCATC 664
 QY 661 AACCTGATCGTCGACGACATCCAGACATTTCTGATGTGACATCTGCAATATGACACCGA 720
 Db 665 AACCTGATCGTCGACGACATCCAGACATTTCTGATGTGACATCTGCAATATGACACCGA 724
 QY 721 GGAGGGTCAATGAGCGGAGCTACCAAGCGGACCTTTTCTGAGCCAGGGGACCACTGGG 780
 Db 725 GGAGGGTCAATGAGCGGAGCTACCAAGCGGACCTTTTCTGAGCCAGGGGACCACTGGG 784
 QY 781 ATGCTGACTCTGCGCAACGGTCACTTTGGAGTTCAGACAGACCCCACTGA 834
 Db 785 ATGCTGACTCTGCGCAACGGTCACTTTGGAGTTCAGACAGACCCCACTGA 838

RESULT 7
 AF125106 1395 bp mRNA linear PRI 07-FEB-2002
 LOCUS Homo sapiens uridine kinase mRNA, complete cds.
 DEFINITION AF125106
 ACCESSION AF125106.1 GI:18568108
 VERSION
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 1395)
 AUTHORS Xiu,Y.R., Yu,L. and Zhao,S.Y.
 TITLE Cloning of a new human cDNA similar to Mus musculus uridine kinase mRNA
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 1395)
 AUTHORS Ding,J.B., Yu,L. and Zhao,S.Y.
 TITLE Direct Submission
 JOURNAL Submitted (02-FEB-1999) Lab of Human Gene Research, Institute of Genetics, Fudan University, No. 220 Handan Rd., Shanghai 200433, People's Republic of China
 FEATURES
 source
 1. 1395
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 8. 838
 /codon_start=1
 /product="uridine kinase"
 /protein_id="AA175943.1"
 /db_xref="GI:18568109"
 /translation="MASAGGAGARAGANRPHQRPPLIGVGGTAGSKSTVEKIME LIGNEVEORHAKVILISDRFYKVLTAEGKAKLKGQYNDHDAEDMDLMTLKN IVEGKTEVPTVDSRLPBTETVVPAADVLEEGILVDSOIRPMPLRFDVD SDVRLSRVRLADVRRGRDLQIILQYTFPKVPAEBSCLPTKRYADYIRGVDNRYA INLIVQHIDILNDICKMKHSGSNRTYKTFSEPDHPQMLTSGKRSHLESSRPH"

BASE COUNT 325 a 352 c 411 g 307 t
 ORIGIN

Query Match 94.2%; Score 786; DB 9; Length 1395;
 Best Local Similarity 97.2%; Pred. No. 6.9e-162;
 Matches 811; Conservative 0; Mismatches 20; Indels 3; Gaps 1;

QY 1 ATGCTTCCGCGGAGCGAGAGCTGCGAGAGCCCGGCGGAGCGGACCGGAACTGACCGGTGT 120
 Db 8 ATGCTTCCGCGGAGCGGAGAGCTGCGAGAGCCCGGCGGAGCGGACCGGAACTGACCGGTGT 64
 QY 61 CAGCGCCCTTCTGATAGGGGTGAGCGCGGCACTGCGACCGGAACTGACCGGTGT 120
 Db 65 CAGCGCCCTTCTGATAGGGGTGAGCGCGGCACTGCGACCGGAACTGACCGGTGT 124
 QY 121 GAGAAATCATGAGATTCTGCGGACAGAAAGAGTGAACGCGGCGGAGGAGGTGTC 180
 Db 125 GAGAAATCATGAGATTCTGCGGACAGAAAGAGTGAACGCGGCGGAGGAGGTGTC 184

QY 181 ATCTGAGCCAGGACAGAGTTCTACAAAGTCTTGACGCGACAGAGGACCAAGCCCTTG 240
 Db 185 ATCTGAGCCAGGACAGAGTTCTACAAAGTCTTGACGCGACAGAGGACCAAGCCCTTG 244
 QY 241 AAAGACAGTACATTTTGCATCCAGATGCGCTTTGATATGATTTGATGACAGACT 300
 Db 245 AAAGACAGTACATTTTGCATCCAGATGCGCTTTGATATGATTTGATGACAGACT 304
 QY 301 CTGAAGAACATCGGTGAGGGGCAAAACGCTGAGAGTCCGACCTATGATTTTGTGACAC 360
 Db 305 CTGAAGAACATCGGTGAGGGGCAAAACGCTGAGAGTCCGACCTATGATTTTGTGACAC 364
 QY 361 TCAAGTTACAGAGACCAAGGTGTCTACCTGCGGACGCTGTTCTGTTGAGGCGATC 420
 Db 365 TCAAGTTACAGAGACCAAGGTGTCTACCTGCGGACGCTGTTCTGTTGAGGCGATC 424
 QY 421 TTGGTGTCTACAGCCAGAGATCCGGAGATGTTCCACTGCGCTCTTCTGTGACACC 480
 Db 425 TTGGTGTCTACAGCCAGAGATCCGGAGATGTTCCACTGCGCTCTTCTGTGACACC 484
 QY 481 GACTCCGACGTCAGGCTGTCTCGAAGATTCGCGGACGTTGGCGGAGGAGGACCTG 540
 Db 485 GACTCCGACGTCAGGCTGTCTCGAAGATTCGCGGACGTTGGCGGAGGAGGACCTG 544
 QY 541 GAGCAGATTCTGACGACGATACCACTTGTGTAAGCGGCGCTTTCGAGAGATTCTGCTG 600
 Db 545 GAGCAGATTCTGACGACGATACCACTTGTGTAAGCGGCGCTTTCGAGAGATTCTGCTG 604
 QY 601 CCGACAAAGAGTATGCCGATGTGATCATCCACAGAGAGTGAACAATATGTTGCCATC 660
 Db 605 CCGACAAAGAGTATGCCGATGTGATCATCCACAGAGAGTGAACAATATGTTGCCATC 664
 QY 661 AACCTGATCGTCGACGACATCCAGACATTTCTGATGTGACATCTGCAATATGACACCGA 720
 Db 665 AACCTGATCGTCGACGACATCCAGACATTTCTGATGTGACATCTGCAATATGACACCGA 724
 QY 721 GGAGGGTCAATGAGCGGAGCTACCAAGCGGACCTTTTCTGAGCCAGGGGACCACTGGG 780
 Db 725 GGAGGGTCAATGAGCGGAGCTACCAAGCGGACCTTTTCTGAGCCAGGGGACCACTGGG 784
 QY 781 ATGCTGACTCTGCGCAACGGTCACTTTGGAGTTCAGACAGACCCCACTGA 834
 Db 785 ATGCTGACTCTGCGCAACGGTCACTTTGGAGTTCAGACAGACCCCACTGA 838

RESULT 8
 AK057848 2228 bp mRNA linear PRI 31-OCT-2001
 LOCUS Homo sapiens cDNA FLJ25119 f1s, clone CBR05878, highly similar to URIDINE KINASE (EC 2.7.1.48).
 DEFINITION AK057848
 ACCESSION AK057848.1 GI:16553809
 VERSION oligo capping, f1s (full insert sequence).
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1
 AUTHORS Tashiro,H., Yamazaki,M., Watanabe,K., Kumagai,A., Itakura,S., Fukuzumi,Y., Fujimori,Y., Komiyaama,M., Suzuki,Y., Hata,H., Nakagawa,K., Mizuno,S., Morinaga,M., Kawamura,M., Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Nishikawa,T., Sugiyama,A., Kawakami,B., Nagai,K., Isogai,T. and Sugano,S.
 TITLE NEDO human cDNA sequencing project
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 2228)
 AUTHORS Sugano,S. and Suzuki,Y.
 TITLE Direct Submission
 JOURNAL Submitted (24-OCT-2001) Sumio Sugano, Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure, Human Genome Center, Shitokane-dai, 4-6-1, Minato-Ku, Tokyo 108-8639, Japan (E-mail:cdna1@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,

COMMENT

Fax:81-3-5449-5416)
 NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction and 5'-end one pass sequencing: Institute of Medical Science, University of Tokyo; Laboratory of Genome Structure, Human Genome Center; 3'-end one pass sequencing: RAB; clone selection for full insert sequencing: RAB and Helix Research Institute.

FEATURES

SOURCE

Location/Qualifiers
 1. .2228
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CBR05878"
 /cissue="brain"
 /clone_id="CBR"
 /note="Cloning vector: PMB18SFL3"

BASE COUNT 478 a 598 c 696 g 456 t
 ORIGIN

Query Match 85.5%; Score 713; DB 9; Length 2228;
 Best Local Similarity 89.2%; Pred. No. 6,7e-146;
 Matches 834; Conservative 0; Mismatches 0; Indels 101; Gaps 2;

1 ATGGCTTCGGCGGGAGCGGAAGACTGCGAGAGCCCGCGCGAGCGCGACCGTCCGAC 60
 64 ATGGCTTCGGCGGGAGCGGAAGACTGCGAGAGCCCGCGCGAGCGCGACCGTCCGAC 123
 61 CAGCGGCGCTTCCTGATAGGGGTGAGCGCGGCACTCCAGCGGGA----- 106
 124 CAGCGGCGCTTCCTGATAGGGGTGAGCGCGGCACTCCAGCGGGAAGCAAGCGGTTG 183
 107 -----AGTCAACCGTGTGTGAAGAGATCATGAGTTG 138
 184 CAGCGGGAATCCCGCTTCTGTCTCTGTCGACGCGCTGTGTGAAGAGATCATGAGTTG 243
 139 CTGGGACAGACGAGGTGAAACAGCGCGAGCGGAAGGTGTCATCTGAGCGAGGACAGG 198
 244 CTGGGACAGACGAGGTGAAACAGCGCGAGCGGAAGGTGTCATCTGAGCGAGGACAGG 303
 199 TTCTACAGGTCCTGACGCGGAGCAGAGGCCCAAGGCTTGAAGAGACAGTCAATTTT 258
 304 TTCTACAGGTCCTGACGCGGAGCAGAGGCCCAAGGCTTGAAGAGACAGTCAATTTT 363
 259 GACCATCAAGTGCCTTTGATGATGATTTGATGACAGACTCTGAAGCAATCTGAGAG 318
 364 GACCATCAAGTGCCTTTGATGATGATTTGATGACAGACTCTGAAGCAATCTGAGAG 423
 319 GGCAGAAACGGTGGAGGTGCGGACCTATGATTTTGTGACACACTCAAGTTACAGAGACC 378
 424 GGCAGAAACGGTGGAGGTGCGGACCTATGATTTTGTGACACACTCAAGTTACAGAGACC 483
 379 ACGGTGCTTACCTGCGGAGCGTGGTCTGTTTGAAGGCACTTGTGTTCTACAGCCAG 438
 484 ACGGTGCTTACCTGCGGAGCGTGGTCTGTTTGAAGGCACTTGTGTTCTACAGCCAG 543
 439 GAGATCGGGAATGTTTCACTGCGGCTTCTGTTGAACACCGACTCTGAGAGCTGAC 498
 544 GAGATCGGGAATGTTTCACTGCGGCTTCTGTTGAACACCGACTCTGAGAGCTGAC 603
 499 TCTGAAAGAGTTCTCGGGAAGTGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 558
 604 TCTGAAAGAGTTCTCGGGAAGTGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 663
 559 TACACCACTTCTGTAAGCCGCGCTTCTGAGAGAGTTCTGCTGCGC----- 603
 664 TACACCACTTCTGTAAGCCGCGCTTCTGAGAGAGTTCTGCTGCGCAGAGCATCTGACAG 723
 604 -----ACAAAGAGATGATCCG 619
 724 GGAATGAGAGTCAAGCATTTGAGCAATGAGTGTGATGAGGAGCAAAAGAGATGATCCG 783
 620 ATGATGATCATCCACGAGAGTGAACAATATGTTGATCAACCTGATGATGATGATGATG 679

Db 784 ATGATGATCATCCACGAGAGGTGAGCAATATGTTTGCATCAACCTGATGTCAGACACA 843
 Qy 680 TCAGAGCATTTGTAATGTTGATCATTTGCAATATGACCCGAGAGAGGATTCATGAGCGGA 739
 Db 844 TCAGAGCATTTGTAATGTTGATCATTTGCAATATGACCCGAGAGAGGATTCATGAGCGGA 903
 Qy 740 GCTACAGCGGAGCCTTTTCTGAGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 799
 Db 904 GCTACAGCGGAGCCTTTTCTGAGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 963
 Qy 800 GGTCAATTTGAGTCCAGAGCAGACCCCACTGA 834
 Db 964 GGTCAATTTGAGTCCAGAGCAGACCCCACTGA 998

RESULT 9
 BC025146
 LOCUS
 DEFINITION Mus musculus uridine monophosphate kinase, mRNA (cDNA clone
 ACCSSION MCC:36231 IMAGE:4913412), complete cds.
 VERSION BC025146
 KEYWORDS MGC.
 SOURCE Mus musculus (house mouse)
 ORGANISM

REFERENCE
 AUTHORS Strausberg, R.L., Feilgold, E.A., Grouse, L.H., Derge, J.G.,
 1 (bases 1 to 1959)
 Strausberg, R.L., Collins, F.S., Wagner, P.H., Schaefer, C.F., Bhat, N.K.,
 Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
 Hopkins, R.F., Jordan, H., Moore, T., Wax, S.I., Wang, D., Hsieh, F.,
 Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
 Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
 Scheetz, T.E., Brownstein, M.J., Uedeln, T.B., Toshiyuki, S.,
 Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
 Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,
 McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
 Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
 Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
 Fahey, J., Helton, E., Ketterman, M., Madan, A., Rodriguez, S.,
 Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
 Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D.,
 Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
 Butlerfield, Y.S., Krzywinski, M.I., Skalski, U., Smalins, D.E.,
 Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

JOURNAL MEDLINE
 PUBMED 12477932
 2 (bases 1 to 1959)
 Strausberg, R.
 Direct Submission
 Submitted (05-MAR-2002) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA

JOURNAL

REMARK

NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
 Contact: MGC help desk
 Email: cgabs-remail.nih.gov
 Tissue Procurement: Jeffrey E. Green, M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Baylor College of Medicine Human Genome
 Sequencing Center
 Center code: BCM-HGSC
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
 Contact: amg@bcm.tmc.edu
 Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Louised, H.,
 Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
 A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAX Plate: 61 Row: P Column: 19
 This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, similarity but not identity to protein.

FEATURES
 source
 1..1959
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="MGC:36231 IMAGE:4913412"
 /issue_type="Salivary gland, 10 week old female mouse"
 /clone_id="Nci CGAP_Sg2"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"

gene

1..1959
 /gene="Umpk"
 /db_xref="locusid:22245"
 /db_xref="MGI:98904"
 57..890
 /codon_start=1
 /product="Umpk protein"
 /protein_id="AAH25146.1"
 /db_xref="GI:19263564"
 /translation="MASAGGSGSAAPEADRPPLIGVSGTSGKSTVCEKTM ELIQNEVDRQRKVLISDQCFKVLAEQKALQGVNFDHPDAFNDLHKITLK NIVGKVEVPTVPTVTHSRLEPTVVPADVLPFBILVEYDEIDMHLRLPVT DSDVRLSRVLRDVRGRDLQILTOYTAFFKPAFEFCPTKXAVIIPRGVDNV AINLVGHIDILNDGLCKRRHGGPNRNMKRTPEPDHGVLTAKRSHLESSSRP H"

CDS

BASE COUNT 466 a 511 c 561 g 421 t
 ORIGIN
 Query Match 78.9%; Score 658; DB 10; Length 1959;
 Best Local Similarity 86.8%; Pred. No. 7.7e-134;
 Matches 724; Conservative 0; Mismatches 110; Indels 0; Gaps 0;

QY 1 ATGGCTTCGGCGGAGGCGGAAGTGCAGAGCCCGCGCGGAGCGGACCTCCGCAC 60
 Db 57 ATGGCTTCGGCGGAGGCGGAGGCGGCGCTCGAGAGCCCGCGCGGAGCGGACCTCCGCAC 116
 QY 61 CAGGGGCGCTTCGATAGGGGTGAGGGGCGGCGCTGCGAGCGGAGTGCACCGTGTGT 120
 Db 117 CCGGGCGGCTTCGATAGGGGTGAGGGGCGGCGGCGCTGAGTGCAGAGTCAACGTGTGT 176
 QY 121 GAGAGATCATGAGTGTCTGGGACAGACGAGGTGGAACAGCGGAGCGGAGGAGTGTCT 180
 Db 177 GAGAGATCATGAGTGTCTGGGACAGACGAGGTGGAACAGCGGAGCGGAGGAGTGTCT 236
 QY 181 ATCTGAGCCAGGAGGCTTCAAGAGTCTGACGCGAGAGCAAGAGCCAGGCTTGG 240
 Db 237 ATCTGAGCCAGGAGGCTTCTCAAGAGTCTGACGCGCGAGAGCAAGAGGCTTGG 296
 QY 241 AAAGAGATGATATTTTGAACATCCAGATCCCTTGAATATGATTTGAGCAAGAGACT 300
 Db 297 AAAGAGATGATATTTTGAACATCCAGATCCCTTGAATATGATTTGAGCAAGAGACT 356
 QY 301 CTGAGAGATCATGAGGCGGCAAAACGCTGAGGCTGCGACCTATGATTTTGTGACAC 360
 Db 357 CTGAGAGATCATGAGGCGGCAAAACGCTGAGGCTGCGACCTATGATTTTGTGACAC 416
 QY 361 TCAAGGTATCCAGAGACCGGTGTCTACCTTCGCGAGCTGTGTTGTTGAGGGATC 420
 Db 417 TCAAGGTATCCAGAGACCGGTGTCTACCTTCGCGAGCTGTGTTGTTGAGGGATC 476
 QY 421 TTGGTGTTCACAGCGAGGAGATCCGGGACATGTTTCAACCTGCGCCCTCTTCGTGGA 480
 Db 477 TTGGTGTTCACAGCGAGGAGATCCGGGACATGTTTCAACCTGCGCCCTCTTCGTGGA 536

QY 481 GACTCCAGCTCAGGCTGTCTCGAAGAGTCTCCGGGACGTGCGCGGAGGAGGACCTG 540
 Db 537 GACTCTGATTTAGGCTGTCTCGAAGAGTCTCCGGGATGTGCAAGAGAGGAGACTG 596
 QY 541 GAGCAGATTTGACGAGTACACACACCTTCTGTAAGCGGCTTTCGAGAGTTCGTGCTG 600
 Db 597 GACGAGATCTGACCTGACGACACCGGCTTGTGAACACGAGCTTTCGAGAGTTCGTGCTG 656
 QY 601 CCGAGCAAGAGATGCGGATGATCATCCGAGAGAGTGTGACATATGTTGCTGATC 660
 Db 657 CCGAGCAAGAGATGCGGATGATCATCCGAGAGAGTGTGATATATGTTGCTGATC 716
 QY 661 AACCTGATCTGACGACATCCAGACATTCGATGATGATGATGATGATGATGATGATG 720
 Db 717 AACCTGATCTGACGACATCCAGACATTCGATGATGATGATGATGATGATGATGATG 776
 QY 721 GAGAGGTTCAGTGGCGGAGCTTCAAGAGGACCTTTTCTGAGCGGAGGACCACTG 780
 Db 777 GCGGGGCGCAACGCGGCGCAACCAAGAGGACCTTCCGAGCGAGATCACCTCGGG 836
 QY 781 ATGCTGACCTTCGCAAGAGCTGACATTCGAGTTCAGAGAGGAGGAGGAGGAGGAG 834
 Db 837 GTGTTGGCAGCTGCAAGGCGCTCACTGAGGTTTACAGAGAGGAGGAGGAGGAGGAG 890

RESULT 10
 BC015547 2072 bp mRNA linear PRI 04-OCT-2001
 Homo sapiens, similar to uridine-cytidine kinase 1, clone MGC:9658
 IMAGE:3845821, mRNA, complete cds.
 BC015547
 BC015547.1 GI:15930229
 MGC.
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 2072)
 Strausberg, R.
 Direct Submission
 Submitted (01-OCT-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 REMARK
 COMMENT
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgabos-remail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: <http://www.shgc.stanford.edu>
 Contact: (Dickson, Mark) mcd@paxil.stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
 R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAX Plate: 20 Row: F Column: 5
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 13899252.

FEATURES
 source

1..2072
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:9658 IMAGE:3845821"
 /issue_type="Colon, adenocarcinoma"
 /clone_id="Nih_MGC_65"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"
 87..692

CDS

```
/codon_start=1
/product="Similar to uridine-cytidine kinase 1"
/product_id="AAH15547.1"
/db_xref="GI:15910230"
/translation="MASAGBDCSPPEADRHOPPLIGVSGTASGKSYCEKIM
ELIGDNVEQRORRVILSDRFYKVLTAQKALQGNFHPDPAFDDLMHRTLK
NIVEGKTEVPTVDVPTVSRILPETVYVPAVDVLFESILVFEYSLEIDMFMHLFVDT
DSPVRLSRDRKEVCRCDHPTRSQYCGHQPDRAAHPHSHSM"
```

BASE COUNT

455 a 559 c 637 g 421 t

ORIGIN

Query Match 76.0%; Score 634; DB 9; Length 2072;
Best Local Similarity 88.6%; Pred. No. 1.4e-128;
Matches 739; Conservative 0; Mismatches 0; Indels 95; Gaps 1;

```
QY 1 ATGGCTTCGGCGGAGCGGAAGACTGGAGAGACCCCGCGCGAGCCGACCTGCGCAC 60
DB 87 ATGGCTTCGGCGGAGCGGAAGACTGGAGAGACCCCGCGCGAGCCGACCTGCGCAC 146
QY 61 CAGCGGCCCTTCTCTAGTAGGGGTGAGCGCGGCACTGCACGGGAAAGTGCACCGTGTGT 120
DB 147 CAGCGGCCCTTCTCTAGTAGGGGTGAGCGCGGCACTGCACGGGAAAGTGCACCGTGTGT 206
QY 121 GAGAAATCATGAGATTCTGGGACAGAACGAGGTGAAACAGCGGACGCGAAGGTGTC 180
DB 207 GAGAAATCATGAGATTCTGGGACAGAACGAGGTGAAACAGCGGACGCGAAGGTGTC 266
QY 181 ATCTGAGCCAGACAGAGTTCTACAGAGTCTTACCGGACAGACGAAAGCCGACCTTG 240
DB 267 ATCTGAGCCAGACAGAGTTCTACAGAGTCTTACCGGACAGACGAAAGCCGACCTTG 326
QY 241 AAAGGACAGTACATTTTGAACATCCAGATGCCCTTATATATGATTGATGACAGGACT 300
DB 327 AAAGGACAGTACATTTTGAACATCCAGATGCCCTTATATATGATTGATGACAGGACT 386
QY 301 CTGAAGAACATCGTGAAGGCAAAACGCTGAGGTGCGACCTATATATTTTGTGACACAC 360
DB 387 CTGAAGAACATCGTGAAGGCAAAACGCTGAGGTGCGACCTATATATTTTGTGACACAC 446
QY 361 TCAAGTTACCAAGACACAGGTGTCTACCTTCGCGACGCTGTTCTGTTGAGGCAATC 420
DB 447 TCAAGTTACCAAGACACAGGTGTCTACCTTCGCGACGCTGTTCTGTTGAGGCAATC 506
QY 421 TTGGTGTCTACAGCCAGGAGATCCGGACATGTTCCACCTGGGCGCTTCGAGGACAC 480
DB 507 TTGGTGTCTACAGCCAGGAGATCCGGACATGTTCCACCTGGGCGCTTCGAGGACAC 566
QY 481 GACTCCGACGTCAAGCTGTCTCGAAGATTCTCCGGACGTGCGCCGAGGGAAGCACTG 540
DB 567 GACTCCGACGTCAAGCTGTCTCGAAGATTCTCCGGACGTGCGCCGAGGGAAGCACTG 593
QY 541 GAGCAGATTCTGACGACGATACACCACTTGTGTAAGCGGCGCTTCGAGGAGTTCTGCTTG 600
DB 594 ----- 593
QY 601 CGGACAAAGAAATATGCGCATGTGATCATCCACGAGAGTGGACAAATATGTTGCCATC 660
DB 594 --GACAAAGAAATATGCGCATGTGATCATCCACGAGAGTGGACAAATATGTTGCCATC 651
QY 661 AACCTGATCTGAGACATCCAGACATTTCTGAATGTGTACATCTTGAATGTGACACCA 720
DB 652 AACCTGATCTGAGACATCCAGACATTTCTGAATGTGTACATCTTGAATGTGACACCA 711
QY 721 GAGAGGTTCATAGGCGGAGCTCAAGCGGACCTTTTCTGACCGAGGGAACCACTCTGG 780
DB 712 GAGAGGTTCATAGGCGGAGCTCAAGCGGACCTTTTCTGACCGAGGGAACCACTCTGG 771
QY 781 ATCTGAGCTCTGTGACAAACGCTCAATTTGAGTCCAGACGAGACCCCACTGA 834
DB 772 ATCTGAGCTCTGTGACAAACGCTCAATTTGAGTCCAGACGAGACCCCACTGA 825
```

RESULT 11

```
BD146824
LOCUS BD146824 753 bp DNA linear PAT 17-JUN-2003
DEFINITION Primer for synthesizing full-length cDNA and use thereof.
ACCESSION BD146824
VERSION BD146824.1 GI:27852582
KEYWORDS JP 2002191363-A/1667.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 753)
Oca,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,D.,
Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
Patent: JP 2002191363-A 1667 09-JUL-2002;
HELIK RESEARCH INSTITUTS
COMMENT
OS Homo sapiens (human)
PN JP 2002191363-A/1667
PD 09-JUL-2002
PF 28-JUL-2000 JP 2000280990
PI TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU
PI JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,
PI KEIICHI NAGAI, TETSUJI OTSUKI
PC C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC
10, C12P21/02, C12Q1/68//C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof FH Key
FEATURES
source
FT source 1..753
FT location/Qualifiers
1..753
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 158 a 196 c 261 g 135 t 3 others
ORIGIN
Query Match 74.8%; Score 623.6; DB 6; Length 753;
Best Local Similarity 97.7%; Pred. No. 2.8e-126;
Matches 642; Conservative 0; Mismatches 12; Indels 3; Gaps 1;
QY 1 ATGGCTTCGGCGGAGCGGAAGACTGGAGAGACCCCGCGAGCCGACCGTCCGCAC 60
DB 95 ATGGCTTCGGCGGAGCGGAAGACTGGAGAGACCCCGCGAGCCGACCGTCCGCAC 154
QY 61 CAGCGGCCCTTCTCTAGTAGGGGTGAGCGGCACTGCACGGGGAAGTGCACCGTGTGT 120
DB 155 CAGCGGCCCTTCTCTAGTAGGGGTGAGCGGCACTGCACGGGGAAGTGCACCGTGTGT 214
QY 121 GAGAAATCATGAGATTCTGGGACAGAACGAGTGGAAACAGCGGACGCGAAGGTGTC 180
DB 215 GAGAAATCATGAGATTCTGGGACAGAACGAGTGGAAACAGCGGACGCGAAGGTGTC 274
QY 181 ATCTGAGCCAGACAGTTCCTACCAAGTCTTGAACGCGAGACAGAAAGCCGACCTTG 240
DB 275 ATCTGAGCCAGACAGTTCCTACCAAGTCTTGAACGCGAGACAGAAAGCCGACCTTG 334
QY 241 AAAGGACAGTACAAATTTTGAACATCCAGATGCCCTTGAATATGATTGATGACAGGACT 300
DB 335 AAAGGACAGTACAAATTTTGAACATCCAGATGCCCTTGAATATGATTGATGACAGGACT 394
QY 301 CTGAAGAACATCGTGAAGGCAAAACGCTGAGGTGCGACCTATATGATTGTGACACAC 360
DB 395 CTGAAGAACATCGTGAAGGCAAAACGCTGAGGTGCGACCTATATGATTGTGACACAC 454
QY 361 TCAAGTTACCAAGACACAGGTGTCTACCTTCGAGACGTGTCTGTTGAGGCAATC 420
DB 455 TCAAGTTACCAAGACACAGGTGTCTACCTTCGAGACGTGTCTGTTGAGGCAATC 514
QY 421 TTGGTGTCTACAGCCAGGAGATCCGGGACATGTTCCACCTGCGGCTCTTCTGAGACAC 480
```


Db 515 TTGGTGTTCACACGACGAGATCCGGACATGTTCCACCTCGCCCTCTTGTGTGACACC 574

Qy 481 GACTCCGACGTGACGCTGTCTCGAAGAGTTCTCCGGAGCTGCCCGGAGGAGGACCTG 540

Db 575 GACTCCGACGTCAGGCTGTCTCGAAGAGTTCTCCGGAGCTGCCCGGAGGAGGACCTG 634

Qy 541 GAGAGATTCTGACGAGTACACACCTTCGTGAAGCGGCTTCGAGGAGTTCTGCC-- 598

Db 635 GAGAGATTCTGACGAGTACACACCTTCGTGAAGCGGCTTCGAGGAGTTCTGCCCT 694

Qy 599 -TGCCGACAAAGATGCGGATGTGATCATCCGACGAGAGTGACAATATGTT 654

Db 695 TGCGACAAAGAGTGTGCGGATGTATCATTCACGAGAGTGACAAATATGANT 751

RESULT 12

MUSURKI 1810 bp mRNA linear ROD 27-MAR-1997

LOCUS Mus musculus uridine kinase mRNA, partial cds.

ACCESSION L31783

VERSION L31783.1 GI:471980

KEYWORDS

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 1810)

AUTHORS Traut,T.W.

TITLE The functions and consensus motifs of nine types of peptide segments that form different types of nucleotide-binding sites

JOURNAL Eur. J. Biochem. 222 (1), 9-19 (1994)

MEDLINE 94259063

PUBMED 8200357

REFERENCE 2 (bases 1 to 1810)

AUTHORS Ropp,P.A. and Traut,T.W.

TITLE Cloning and expression of a cDNA encoding uridine kinase from mouse brain

JOURNAL Arch. Biochem. Biophys. 336 (1), 105-112 (1996)

MEDLINE 97108719

PUBMED 8951040

FEATURES

source location/Qualifiers

1..1810 /organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

/tissue_type="Brain"

<1..783 /EC_number="2.7.1.48"

/standard_name="uridine-cytidine kinase"

/codon_start=1

/product="uridine kinase"

/protein_id="AAB50568.1"

/db_xref="GI:471981"

/translation="RPPRPFLIGVSGTASGKSTVEKIMELIGONEVDRORKLVI LSQCFKVLTAQKAKLQGVNFDHPDAMDLMKTKLKNIVEGTVAPVPTDFT HSRLEPTVVPPADIVLFEGLIVYQOEDIMELRLPVTDSVRLSRVLDVQNG RDLBQILTYTAFYKPAFBEFCLEPTKXADVIIRGVNDVAILVQHTQDILNGDL CKHRGGPNGNHNRKTPPEPDHGVATGRSHLESSRPH"

misc_binding join(37..63,412..426,451..465)

misc_binding /bound_molety="ATP at catalytic site"

BASE COUNT 1790 a 482 c 512 g 405 t

ORIGIN

Query Match 74.3%; Score 619.8; DB 10; Length 1810;

Best Local Similarity 87.0%; Pred.No.1.8e-125;

Matches 681; Conservative 0; Mismatches 102; Indels 0; Gaps 0;

Qy 52 CGTCCGACCGGCGCTTCTCTATAGGGGTGACGGCGGACATGCCAGCGGGAAGTCG 111

Db 1 CGTCCGACGCGCGCGCTTCTCTATAGGGGTGACGGCGGACCGGACCGGTAAGTCA 60

Qy 112 ACCGTGTGGAAGATCATGAGTTGCTGGACAGAACGAGTGGACAGCGGACGG 171

Db 61 ACAGTGTGGAAGATCATGAGTTGCTGGACAGAACGAGTGGACAGCGGACGG 120

Qy 172 AAGTGTGATCTCTGAGCGAGACTCTTCTCAAGATTCTGACGCGGAGGAGGCC 231

Db 121 AAGTGTGATCTCTGAGCGAGACTCTTCTCAAGATTCTGACGCGGAGGAGGCC 180

Qy 232 AAGCTCTGGAAGAGACATATTTTGAACATCCAGATGCTTGTATATATTTGAG 291

Db 181 AAGCTCTGGAAGAGACATATTTTGAACATCCAGATGCTTGTATATATTTGAG 240

Qy 292 CACAGACTCTGAAGAATCGTGAAGGCAAAACGGTGAAGTGGCGACTGATTTT 351

Db 241 CACAGACTCTGAAGAATCGTGAAGGCAAAACGGTGAAGTGGCGACTGATTTT 300

Qy 352 GTGACACTCAAGTTACCGAGACACGCTGTACCTCGGACGCTGTCTTT 411

Db 301 GTGACACTCAAGTTACCGAGACACGCTGTACCTCGGACGCTGTCTTT 360

Qy 412 GAGGCACTGTGTGTTTACAGTCAAGAGATCGGACATGTTCCACTGCTCTTTC 471

Db 361 GAGGCACTGTGTGTTTACAGTCAAGAGATCGGACATGTTCCACTGCTCTTTC 420

Qy 472 GTGACACTCAAGTTACCGAGACACGCTGTCTCGAAGATTCTCCGGACGCTGAGG 531

Db 421 GTGACACTCAAGTTACCGAGACACGCTGTCTCGAAGATTCTCCGGACGCTGAGG 480

Qy 532 AGGACCTGAGAGATTTGACGAGTACACCACTTGTGAAGCGGCTTTGAGAG 591

Db 481 AGGACCTGAGAGATTTGACGAGTACACCACTTGTGAAGCGGCTTTGAGAG 540

Qy 592 TTCTGCTGCGGCAAAAGATGCGGATGTATATATCCACAGAGATGGAATATG 651

Db 541 TTCTGCTGCGGCAAAAGATGCGGATGTATATATCCACAGAGATGGAATATG 600

Qy 652 GTTGACATCAACTGATCGGACGACATTCAGAGATTTGATGTGATCTGCAAA 711

Db 601 GTTGACATCAACTGATCGGACGACATTCAGAGATTTGATGTGATCTGCAAA 660

Qy 712 TGGCACCAGAGAGTTCATGAGCGGAGCTACAGAGACCTTTTCTGAGCGAGGAC 771

Db 661 TGGCACCAGAGAGTTCATGAGCGGAGCTACAGAGACCTTTTCTGAGCGAGGAC 720

Qy 772 CACCTGAGATGCTGACCTCTGCAAGAGTCACTTTGAGATGACAGACCCAC 831

Db 721 CACCTGAGATGCTGACCTCTGCAAGAGTCACTTTGAGATGACAGACCCAC 780

Qy 832 TGA 834

Db 781 TGA 783

RESULT 13

MUSURKI 952 bp DNA linear ROD 05-MAR-2001

LOCUS Mus musculus uridine kinase pseudogene.

DEFINITION L31784.1 GI:471982

ACCESSION L31784.1

VERSION L31784.1

KEYWORDS pseudogene; uridine kinase.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 952)

AUTHORS Traut,T.W.

TITLE The functions and consensus motifs of nine types of peptide segments that form different types of nucleotide-binding sites

JOURNAL Eur. J. Biochem. 222 (1), 9-19 (1994)

MEDLINE 94259063

PUBMED 8200357

REFERENCE 2 (bases 1 to 952)

AUTHORS Ropp,P.A. and Traut,T.W.

TITLE Cloning and expression of a cDNA encoding uridine kinase from mouse brain
 JOURNAL Arch. Biochem. Biophys. 336 (1), 105-112 (1996)
 MEDLINE 97108719
 PUBMED 8951040
 FEATURES Location/Qualifiers
 source 1..952
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /tissue_type="kidney"
 76..913
 /standard_name="uridine-cytidine kinase"
 /pseudo
 /codon_start=1
 /product="uridine kinase"
 misc_difference 130
 /note="insertion not found in L31783"
 BASE COUNT 230 a 231 c 277 g 214 t
 ORIGIN

Query Match 70.9%; Score 591.2; DB 10; Length 952;
 Best Local Similarity 82.5%; Pred. No. 3.5e-119;
 Matches 691; Conservative 0; Mismatches 143; Indels 4; Gaps 1;

1 ATGGCTTCGGCGGAGGCGAAGACTGGAGCCCGC-----GCCGGAGGCCGACCGCTCC 56
 76 ATGGCTTCGGCGGAGGCGGCTGCTCAGAGAGTCTGCGCTGCCGCGGCGCATCTTCC 135
 57 GCACCAAGCGCCCTTCTGATAGGGGTGAGCGGCGGCACTGCCAGCGGAAGTCAACCT 116
 136 TCAGCCCTTGGCCGTTCTCATCGGCGGTGAGGGGCGGACCGCTAGTGGCAAGTCAACGT 195
 117 GTGTGAGAAATCATGAGATTGCTGGGACAGAACGAGCTGGAACAGCGGACGGAAAGT 176
 196 GTGTGAGAAATCATGAGCTGCTGGGACAGAAATGAAGTGAACCTCCGCTAGCGCAAGTT 255
 177 GGTGATCTGAGCGAGGACAGGTTCTCAAGGTCTGACGGGACAGAGCAAGGCAAGGCAAGG 236
 256 GGTGATCTGAGCGAGGACAGGCTCTCAAGGTCTGACGGGACAGAGCAAGGCAAGGCAAGG 315
 237 CTTGAAAGGACAGTACATTTTTCACATCCAGATGCTTTGATATGATTTGATGACAG 296
 316 TTGGAAGGACAGTACATTTTTCACATCCAGATGCTTTGATATGATTTGATGACAG 375
 297 GACTCTGAGAAACATCTGTGAGAGGCAAAACGCTGAGAGTCCGACCTATGATTTGTGAC 356
 376 GACTCTGAGAAACATTTTGAAGGCAAAAGTTTAAAGTTCTTACCTATGATTTGTGAC 435
 357 ACATCAAGGTTACGAGACCAACGCTGCTACCCCTCGGACGCTGCTTGTGAGAGG 416
 436 CGACTCAAGGTTACGAGTACCACTGTGCTACCAATTTGATGAGCGGTTCTGTTCAGGG 495
 417 CATCTTGTTGTTTACAGCAGGAGATCCGGGACATGTTCCACTCGGCTTCTGCGA 476
 496 CATCTTGTTGTTTACAGCAGGATCCGGGACATGTTCCACTCGGCTTCTGCGA 555
 477 CACCGACTCCGAGCTCAGGCTGTCTGAGAGATTTCTCGGACGCTGCGGAGGAGGA 536
 556 CACAGACCCCTGATGTTAGGCTGTCTTGAAGAGTTCTCGGGAATGTGCMAAGAGGGA 615
 537 CCGGAGCAGATTCTGACGAGTACACCACTTCGTAAGCCGCGCTTCGAGAGTTCTG 596
 616 CCGGAGCAGATTCTGACGAGTACACCACTTCGTAAGCCGCGCTTCGAGAGTTCTG 675
 597 CCGGAGCAGAAAGATATCCGATGATCATCCAGAGAGTGAACATATGTTGC 656
 676 CCGGAGCAGTAAAGATAGCTGACGATCATCCCTCGAGGGCTTGACAAATATGTTGC 735
 657 CATCAACCTATGTTGAGACATCCAGGACATTTGAATGTTGACATCTGCAATGGCA 716
 736 CATCAACCTGATGTTATACATCCAGGACATCTTCACGAGGACCTGTGCAAGCGGCA 795

Qy 717 CCAGAGAGGATCCATATGAGCGGAGTACAGCGGACCTTTTCTGAGCCAGAGGACCC 776
 Db 796 CCGAGCGGCGCCCAATGGGCGACAGCCCAAGAGACTTATCCGATCTCAGATCATCCC 855

Qy 777 TGGGATGCTGACCTCTGGCAAAACGCTACATTTGAGTCCAGACAGACCCCATGA 834
 Db 856 TGGGATGTTGGCACATGTTAAACGCTGACACCTGAGTCTAGACAGACAGACCCATGA 913

RESULT 14
 AC078885/c
 LOCUS
 DEFINITION
 Mus musculus chromosome 6 clone RP23-117123 strain C57BL6/J,
 WORKING DRAFT SEQUENCE, 44 unordered pieces.
 AC078885
 AC078885.8 GI:18376842
 HTG; HTGS PHASE1; HTGS DRAFT.
 Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 254993)
 Montgomerly,K.T., Grills,G., Han,J., Lee,E., Long,J., Pomerantz,R.,
 Ioshikhes,I.P., Shim,C., Decker,J., Thomas,E., Pereira,A.,
 Gordon,M., Goltz,J.S. and Kucherlapati,R.
 High Throughput Mouse Sequencing
 Unpublished
 2 (bases 1 to 254993)
 Montgomerly,K.T., Grills,G., Han,J., Lee,E., Long,J., Pomerantz,R.,
 Ioshikhes,I.P., Shim,C., Decker,J., Thomas,E., Pereira,A.,
 Gordon,M., Goltz,J.S. and Kucherlapati,R.
 Direct Submission
 Submitted (08-AUG-2000) Department of Molecular Genetics, Albert
 Einstein College of Medicine Genome Center, 1300 Morris Park Ave.,
 Bronx, NY 10461, USA
 On Jan 26, 2002 this sequence version replaced gi:14488282.
 -----Genome Center
 Center: Harvard Partners Genome Center
 Center Code: HPGC
 Web site: http://www.hpcgg.org/Sequence/mouse.html
 Contact: hpcg@mednet.mgh.harvard.edu
 -----Summary Statistics
 Center project name: AAP
 Sequencing vector: pUC18; L08752
 Chemistry: Dye-terminator Big Dye; 100%
 *Consensus quality: 247096 at least Q20
 *Consensus quality: 242518 at least Q30
 *Consensus quality: 235589 at least Q40
 Estimated insert size: agarose-PP - N/A
 **Estimated insert size: 254133 - sum-of-contigs
 Quality coverage: agarose-PP - N/A
 Quality coverage: 6 x in Q20 bases; sum-of-contigs estimation

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 44 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 27415: contig of 27415 bp in length
 * 27416: gap of unknown length
 * 27436 57817: contig of 30382 bp in length
 * 57818 57837: gap of unknown length
 * 57838 81637: contig of 23800 bp in length
 * 81638 81657: gap of unknown length
 * 81658 102155: contig of 20498 bp in length
 * 102156 102175: gap of unknown length
 * 102176 121200: contig of 19025 bp in length
 * 121201 121220: gap of unknown length
 * 121221 135073: contig of 13853 bp in length
 * 135074 135093: gap of unknown length

135034	156689:	contig of 21596 bp in length
156690	156705:	gap of unknown length
156710	156785:	contig of 9476 bp in length
166206	166205:	gap of unknown length
166266	178798:	contig of 12593 bp in length
178819	178818:	gap of unknown length
178847	187247:	contig of 8429 bp in length
187248	187267:	gap of unknown length
187268	195492:	contig of 5225 bp in length
192493	195512:	gap of unknown length
192595	212964:	gap of unknown length
212965	219468:	contig of 6504 bp in length
219469	219488:	gap of unknown length
219489	220369:	contig of 881 bp in length
220370	220389:	gap of unknown length
220390	221164:	contig of 775 bp in length
221165	221184:	gap of unknown length
221185	222796:	contig of 1612 bp in length
222797	222816:	gap of unknown length
222817	224762:	contig of 1946 bp in length
224763	224782:	gap of unknown length
224783	225999:	contig of 1217 bp in length
226000	226019:	gap of unknown length
226001	227409:	contig of 1390 bp in length
227410	227429:	gap of unknown length
227430	230145:	contig of 2716 bp in length
230146	230165:	gap of unknown length
230166	231842:	contig of 1677 bp in length
231843	231862:	gap of unknown length
231863	237786:	contig of 924 bp in length
232787	232806:	gap of unknown length
232807	233137:	contig of 331 bp in length
233138	233157:	gap of unknown length
233158	234519:	contig of 1362 bp in length
234520	234539:	gap of unknown length
234540	236416:	contig of 1677 bp in length
236417	236436:	gap of unknown length
236437	237967:	contig of 1531 bp in length
237968	239320:	contig of 1233 bp in length
239221	239240:	gap of unknown length
239241	239267:	contig of 127 bp in length
239268	239387:	gap of unknown length
239388	239399:	contig of 572 bp in length
239399	239599:	gap of unknown length
240107	240161:	contig of 1037 bp in length
241017	241036:	gap of unknown length
241037	242391:	contig of 1355 bp in length
242392	244111:	gap of unknown length
244112	244823:	contig of 1412 bp in length
244824	244843:	gap of unknown length
244844	246227:	contig of 1384 bp in length
246228	246247:	gap of unknown length
246248	246334:	contig of 1087 bp in length
246335	246354:	gap of unknown length
246355	247451:	contig of 1097 bp in length
247452	247471:	gap of unknown length
247472	248780:	contig of 1309 bp in length
248781	248800:	gap of unknown length
248801	250041:	contig of 1241 bp in length
250042	250061:	gap of unknown length
250062	251173:	contig of 1112 bp in length
251174	251193:	gap of unknown length
251194	255733:	contig of 1520 bp in length
252714	252733:	gap of unknown length
252734	254014:	contig of 1281 bp in length
254015	254034:	gap of unknown length
254035	255176:	contig of 142 bp in length

[illegible]

QY 115 GTGTTGAGAAGATCATGAGTTGCTGGGACAGAACGAGGTGAAACGCGGACGCGAG 174
 Db 156555 GTGTGTAGAAAGATCAGAGAGCTCTGTGGAGAGAACGAGTGGACCGCGGACGCGAG 156596
 QY 175 GTGTGATCTCTGACCGAGACGAGTTGTACAAAGGTTCTCTGACGCGAGACGAGAACGCCAAG 234
 Db 156595 TTGGTATCTCTGACCGAGAGTTGCTTACAAAGTTCTGACGCGGACGAGAACGCCAAG 156536
 QY 235 GCCTTGAAGAGACGTCATATTTGACCATTCAGATGCTTTGATTAATGATTTGATGCAC 294
 Db 156535 GCCTTGAAGAGACGTCATATTTGACCATTCAGATGCTTTGATTAATGATTTGATGCAC 156476
 QY 295 AGGACTGTGAAGACATCTGTGAGAGGAAAAACGGTGGAGGCGGACCTATGATTTGTG 354
 Db 156475 AAGACCTGTGAAGACATTTGTGAAGGAAAACTGTGAGGTCCTTACCTATGATTTGTG 156416
 QY 355 ACACTCAAGGTATCCAGAGACGAGTGTGTCTACCTCGGACGTCGTTCTTTGTGAG 414
 Db 156415 ACCACTCAAGGTATCCAGAGACGAGTGTGTCTACCTCGGACGTCGTTCTTTGTGAG 156356
 QY 415 GGCACTTTGGTGTCTTCAAGCAGAGAAATCCGGACATGTTCCACTGCGCTCTTCTGTG 474
 Db 156355 GGCACTTTGGTGTCTTCAAGCAGAGAAATCCGGACATGTTCCACTGCGCTCTTCTGTG 156296
 QY 475 GACACGAGCTCCGACGTCAGGCTGTCTCGAAGAGTTCTCCGGGACGTCGCGGAGGAGG 534
 Db 156295 GACACGAGCTCTGTATGTTAGGCTGTCTCGAAGAGTTCTCCGGGACGTCGCGGAGGAGG 156236
 QY 535 GACCTGAGAGATTTCTGACGAGTACGACCTTCTCGTGAAGCGCGCTTCTGAGAGATTTC 594
 Db 156235 GACCTGAGAGATTTCTGACGAGTACGACCTTCTCGTGAAGCGCGCTTCTGAGAGATTTC 156176
 QY 595 TGCTTCCGACAAAGATGTCGATGTGATTCACGAGAGTGTGACATATGTTG 654
 Db 156175 TGCTTCCGACAAAGATGTCGATGTGATTCACGAGAGTGTGATATATGTTG 156116
 QY 655 GGCATCAACCTGATCGAGACATCCAGAGACCTTGTGAGAGAGGAGACAC 714
 Db 156115 GGCATCAACCTGATCGAGACATCCAGAGACCTTGTGAGAGAGGAGACAC 156056
 QY 715 CACGAGAGAGGTCATATGAGCGAGTACAGACGACCTTTTGTGAGCGAGGAGACAC 774
 Db 156055 CACGAGAGAGGTCATATGAGCGAGTACAGACGACCTTTTGTGAGCGAGGAGACAC 155996
 QY 775 CTTGGAGTCTGACCTTGTGAAACGTCACATTTGAGTCCAGACGAGACCCACTGA 834
 Db 155995 CTTGGAGTCTGACCTTGTGAAACGTCACATTTGAGTCCAGACGAGTCCAGACCCACTGA 155936

RESULT 15
 AC118474 192336 bp DNA linear HTG 25-MAY-2003
 AC118474 Mus musculus clone RP23-288012, *** SEQUENCING IN PROGRESS ***, 4
 LOCUS unordered pieces.
 DEFINITION
 AC118474 GI:31074905
 AC118474 HTG; HTGS_PHASE1; HTGS_FUILLTOP; HTGS_ACTIVEFIN.
 KEYWORDS Mus musculus (house mouse)
 SOURCE Mus musculus
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eumetazoa; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 192336)
 2 (bases 1 to 192336)
 Mus musculus, clone RP23-288012
 Unpublished
 Birren,B., Linton,L., Nussbaum,C., Lander,E., Ali,A., Allen,N.,
 Anderson,S., Barina,N., Bastien,V., Bloom,T., Boguslavsky,L.,
 Boukhalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
 Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,
 Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
 Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,

Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
 Hago,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
 Kamat,A., Karatas,A., Kells,C., Labocque,K., Lamazates,R.,
 Landers,T., Lenoczky,J., Levine,R., Lindblad-Toh,K., Liu,G.,
 Maclean,C., MacDonald,P., Major,J., Margulis,N., Matthews,C.,
 McCarthy,M., McEwan,P., McKernan,K., Meldrum,J., Meunier,L.,
 Mihova,T., Mlenka,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,
 Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
 Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
 Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
 Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schuback,R.,
 Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
 Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
 Topham,K., Travers,M., Travis,N., Trisillo,J., Vassiliev,H.,
 Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
 Zainoun,J., Zembek,L., Zimmer,A., and Zody,M.
 Direct Submission
 Submitted (17-APR-2002) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 3 (bases 1 to 192336)

REFERENCE
 AUTHORS
 Birren,B., Nussbaum,C., Lander,E., Abouelleil,A., Allen,N.,
 Anderson,S., Arachchi,H.M., Barina,N., Bastien,V., Bloom,T.,
 Boguslavsky,L., Boukhalter,B., Camarata,J., Chang,J., Choepel,Y.,
 Collymore,A., Cook,A., Cooke,P., Cornu,B., Dearellano,K.,
 Diaz,J.S., Dodge,S., Doolley,K., Dorris,L., Erickson,J., Faro,S.,
 Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J., Gardyna,S.,
 Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hago,B.,
 Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
 Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
 Lindblad-Toh,K., Liu,G., Lui,A., Mabbitt,R., Maclean,C.,
 MacDonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,
 Meldrum,J., Meunier,L., Mihova,T., Mlenka,V., Murphy,T., Naylor,J.,
 Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
 O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
 Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
 Roman,J., Schauer,S., Schuback,R., Seaman,S., Severy,P., Smith,C.,
 Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,
 Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M.,
 Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
 Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A., and Zody,M.
 Direct Submission
 Submitted (25-MAY-2003) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On May 25, 2003 this sequence version replaced gi:28975887.
 All repeats were identified using RepeatMasker:
 Smit,A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RW/RepeatMasker.html

Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu
 Project Information
 Center project name: L23680
 Center clone name: 288_O_12

NOTE: This is a 'working draft' sequence. It currently
 consists of 4 contigs. The true order of the pieces
 is not known and their order in this sequence record is
 arbitrary. Gaps between the contigs are represented as
 runs of N, but the exact sizes of the gaps are unknown.
 This record will be updated with the finished sequence
 as soon as it is available and the accession number will
 be preserved.
 1 36631: contig of 36631 bp in length
 * 36632 36731: gap of 100 bp
 * 36732 160664: contig of 123933 bp in length
 * 160665 160764: gap of 100 bp
 * 160765 177955: contig of 17191 bp in length
 * 177956 178055: gap of 100 bp
 * 178056 192336: contig of 14281 bp in length.
 Location/Qualifiers
 1. 192336

FEATURES
 source

```

/organism="Mus musculus"
/mol_type="Genomic DNA"
/db_xref="taxon:10090"
/clone="RP23-28012"
/clone_lib="R1C1-23 Female Mouse BAC"
BASE COUNT 58391 a 36846 c 37829 g 58262 t 1008 others
ORIGIN

```

	Query Match	67.9%;	Score 566;	DB 2;	Length 192336;
	Best Local Similarity	81.6%;	Pred. No. 8,9e-114;		
	Matches 684;	Conservative 0;	Mismatches 140;	Indels 14;	Gaps 2;
QY	1	ATGCTTCGCGCGGAGCGGAAGACTGCGAGACCCCGC---	-CGCGAGGCCGACCGTCC	56	
Db	77706	ATGCTTCGCGCGGAGCGGCTGCTCGAGAGTACTGCGCCCTCCCGAGGCGCATCTTCC		777655	
QY	57	GCACGAGCGGCCCTTCCTGATAGGGGAGACCGCGGACCTGCGAGGGGAAAGTCACCGT		116	
Db	77766	TCAAGCTTGGCCGCTTCCTCATCGCGGAGTGGGGGACCGCTAGTGGCAAGTCAACAT		778235	
QY	117	GTTGAGAAAGATCATGAGATTGCTGGGACAAACGAGGTGAAACAGCGCAGCGGAAGT		176	
Db	77826	GTTGAGAAAGATCATGAGCTGCTGGGACGAATTAAGTGAACCTCGGAGGCAAGTT		778855	
QY	177	GGTATCTCTGAGCCAGGACAGGTTCTACAGGTCCTGACGGCAGACAGAACGCCAAAGC		236	
Db	77886	GGTATCTCTGAGCCAGGACGCTCTTCAAGAGTTCTGACCGCTGACCAAGAACCCAAAGC		779455	
QY	237	CTTGAAAGGACAGTACATAATTTTGACATCCAGATGCTTGTGATATGATTTGATGACAG		296	
Db	77946	TTTGAAGGACAGTACAAATTTTGACATGAGATGCTTTGATATGATCTGATGACAA		780055	
QY	297	GACTCTGAGAACATCGTGAAGGGCAAAACGTTGAGGTCCGACTATGATTTGTGTGAC		356	
Db	78006	GACTCTGAAAAACATTTTAAAGGCAAAAGTTGTTAAGTTCCTAACCTAATGATTTGTGTGAC		780655	
QY	357	ACACTCAAGGTTTACCAAGACCAACGGGTCATCCCTCGCGGACGTTGTTGAGGG		416	
Db	78066	CACTCAAGGTTTACCAAGTACCACTGTGTGTCACAAATTAATGATCGGTTCTGTTCGAGGG		781255	
QY	417	CATCTGTGTGTTCTACAGCCAGAGATCCGGAACATGTTCCACTCGGCTCTTCGTGTA		476	
Db	78126	CATCTGTGTATTTCTACACTCAAGAGATCCGGGACATGTTTCACTCGGCTCTTCGTGTA		781855	
QY	477	CACCGACTCCGACGTCAGGCTGTCTCGAAGATTTCCCGGACGTTCCCGGAGGAGGGA		536	
Db	78186	CACGACCCCTGAGTATGAGCTGTCTGTAAGATTTCTGGGATGTCAATGAGGAAGGGA		782455	
QY	537	CTTGAGACAGATTCTGACGAGTACACACCTTCGTGAAACCGGCTTCGAGAGAGTTCTG		596	
Db	78246	CTTGAGACAGATCTCTGACTCAGTGACACACATTTGTGAATATGGGCTTCGAGAGAGTTCTG		783055	
QY	597	CTGCGCAGCAAAAGAGATGTCGCAATGTGATCATCCACGAGAGTGGACAAATATGGTTGC		656	
Db	78306	CTGCGCAGCTTGAAGAA-----GTGATCATCTCTCGAGGGCTTGAACAAATATGGTTGC		783555	
QY	657	CATCAACCTATCGTGCAGACATCCAGAGCATTTCTGAATGTGACATCTGCAAAATGGGA		716	
Db	78356	CATCAACCTATCGTATTAACATCATCCAGAGCATCTCTCAACGGGACCTGTGCAAGCGGCA		784155	
QY	717	CCGAGAGGGTCCATGAGGCGGAGCTAACAGCGGACCTTTTCTGAGCAAGGGGACCAACC		776	
Db	78416	CCGAGAGGGGCCCATGAGGCGACAGCAACAGAGGACCTTAACCGAGTCAAGAGATCAACC		784755	
QY	777	TGGGATCTCTGACCTCTGGGCAAAAGGTCACTTTGAGTCTCAGAGAGACCCCATCTGA		834	
Db	78476	TGGGATTTTGGCCATCTGTATTAAGCTGTACACCTGTGAGTCTTGAAGAGAGATCTCCATTGA		78533	

Search completed: November 25, 2003, 00:22:05
Job time : 3247.66 secs

PR	19-OCT-2000; 2000US-0693036.
PR	29-NOV-2000; 2000US-0727344.
XX	
PA	(HYSE-) HYSEQ INC.
XX	
PI	Tang YT, Liu C, Aesundl V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI	Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI	Zhao QA, Zhou P, Goodrich R, Dymnac RT;
XX	
DR	WPI: 2001-442253/47.
DR	P-PSDB; AAM39502.
XX	
PT	Novel nucleic acids and polypeptides, useful for treating disorders
PT	such as central nervous system injuries -
XX	
PS	Claim 1, SEQ ID NO 861; 10078p; English.
XX	
CC	The invention relates to human nucleic acids (AA157798-AA161369) and
CC	the encoded polypeptides (AAM38642-AAW42213) with nootropic,
CC	immunosuppressant and cytostatic activity. The polynucleotides are useful
CC	in gene therapy. A composition containing a polypeptide or polynucleotide
CC	of the invention may be used to treat diseases of the peripheral nervous
CC	system, such as peripheral nervous injuries, peripheral neuropathy and
CC	localised neuropathies and central nervous system diseases, such as
CC	Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC	lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC	utilisation of the activities such as: Immune system suppression,
CC	Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC	and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC	assays for receptor activity, arthritis and inflammation, leukaemias and
CC	C.N.S disorders.
CC	Note: The sequence data for this patent did not form part of the printed
CC	specification.

Seq	Sequence	1288 Bp	275 A	361 C	418 G	234 T	0 other
Query Match	100.0%	Score 834	DB 22	Length 1288			
Best Local Similarity	100.0%	Pred. No. 0					
Matches	834	Conservative	0	Mismatches	0	Indels	0
						Gaps	0
QY	1	ATGGCTTCGGCGGGAGCGGAAGCTCGAGACGCCCGCGCGCGAGAGCCGACCGTCGCGAC	60				
Db	95	ATGGCTTCGGCGGGAGCGGAAGCTCGAGAGCCCGCGCGCGAGAGCCGACCGTCGCGAC	154				
QY	61	CAGCGGCCCTTCCTGATATAGGGGTGAGCGGCGCACTGCCAGCGGAAAGTCGACCGTGTGT	120				
Db	155	CAGCGGCCCTTCCTATATAGGGGTGAGCGGCGCACTGCCAGCGGAAAGTCGACCGTGTGT	214				
QY	121	GAGAAAGATCATGAGAAATTCGTGGGACAGAAACGAGGTGTGAAACACCGCAGCGGAAAGTGTCT	180				
Db	215	GAGAAAGATCATGAGAAATTCGTGGGACAGAAACGAGGTGTGAAACACCGCAGCGGAAAGTGTCT	274				
QY	181	ATCTGAGCCACGAGACAGGTTCTACAAAGTCTCAGCGGACAGACAGAAAGCCAAAGGCGTTG	240				
Db	275	ATCTGAGCCACGAGACAGGTTCTACAAAGTCTCAGCGGACAGACAGAAAGCCAAAGGCGTTG	334				
QY	241	AAAGGACAGTACAAATTTTGAACCATTCAGATGCTTTGATATGATTTGATGACAGAGACT	300				
Db	335	AAAGGACAGTACAAATTTTGAACCATTCAGATGCTTTGATATGATTTGATGACAGAGACT	394				
QY	301	CTGAAGAACATCTGTGGAGGGCAAAACGGTGGAGGTGCGGACCTATGATTTTGTGACACAC	360				
Db	395	CTGAAGAACATCTGTGGAGGGCAAAACGGTGGAGGTGCGGACCTATGATTTTGTGACACAC	454				
QY	361	TCAAAGTTACAGAGAACCAACGGTGTGTTACCTCTGGAGACGTGTGTTCTGTTTAAAGGCGATC	420				
Db	455	TCAAAGTTACAGAGAACCAACGGTGTGTTACCTCTGGAGACGTGTGTTCTGTTTAAAGGCGATC	514				
QY	421	TTGTGTTCTACAGCCAGAGAGATCCGGGACATGTTCCACTCGGCTCTTCGTGACACC	480				
Db	515	TTGTGTTCTACAGCCAGAGAGATCCGGGACATGTTCCACTCGGCTCTTCGTGACACC	574				
QY	481	GACTTCGACGTACGGCTGTCTTCGAAAGATTCTCCGGGACGTGCGCCGAGGAGGACCTTG	540				

Db	575	GACTCCGACCGTCAGGGCTGTCTCCGAAGATTCTCCGGGACGTGCGCCGAGGAGGAACTTG	634
Oy	541	GAGCAGATTCTGACGCGATACACCACCTTCTGTGAAGCCGGCCCTTCCAGGAGTTCTGCTTG	600
Db	635	GAGCAGATTCTGACGCGATACACCACCTTCTGTGAAGCCGGCCCTTCCAGGAGTTCTGCTTG	694
Oy	601	CCGCAAAAGAAGTATGCGCATGTGATCATCCCAAGAGATGGACAATATGTTGCCATC	660
Db	695	CCGCAAAAGAAGTATGCGCATGTGATCATCCCAAGAGATGGACAATATGTTGCCATC	754
Oy	661	AACTGATCGTGACGACATCCAGGACATTTCTGATATGTGACATCTGCAATATGGCACCGA	720
Db	755	AACTGATCGTGACGACATCCAGGACATTTCTGATATGTGACATCTGCAATATGGCACCGA	814
Oy	721	GGAGGATCCAAATGGGGCGAGCTCAAGCGGACCTTTTCTGAGCCAGGGGACCAACCTTG	780
Db	815	GGAGGATCCAAATGGGGCGAGCTCAAGCGGACCTTTTCTGAGCCAGGGGACCAACCTTG	874
Oy	781	ATGCTGACCTCTGGCAAAACGGTCACATTTTGAAGTCCAGCAGACAGACCCCACTGA	834
Db	875	ATGCTGACCTCTGGCAAAACGGTCACATTTTGAAGTCCAGCAGACAGACCCCACTGA	928
RESULT 2			
ID	AAD27186	standard; DNA; 1624 BP.	
AC	AAD27186;		
DT	18-APR-2002	(first entry)	
DE	Human 57658 DNA.		
XX	Human; uridine kinase-like protein; haematopoietic neoplastic disorder;		
KW	57658 protein; leukaemia; haematopoietic disorder; diabetes mellitus;		
KW	rheumatoid arthritis; psoriasis; dermatitis; osteoporosis; rickets;		
KW	barcoma; myocardial infarction; hypertension; atherosclerosis;		
KW	Alzheimer's disease; Parkinson's disease; Huntington's disease;		
KW	gene therapy; chromosome mapping; tissue typing; dermatological;		
KW	cytostatic; osteopathic; cardiac; neuroprotective; nootropic;		
KW	anticonvulsant; ds.		
XX			
OS	Homo sapiens.		
XX			
XX	Key	Location/Qualifiers	
XX	CDS	94..927	
FT	FT	/*tag= a	
FT	FT	/product= "Human 57658 protein"	
FT	FT	94..924	
FT	FT	/*tag= b	
FT	FT	/note= "this region is specifically claimed as	
FT	FT	SEQ ID NO: 3 in claim 1 of the specification"	
XX			
XX	MO200202761-A2.		
XX			
XX	10-JAN-2002.		
XX			
XX	28-JUN-2001; 2001MO-US21063.		
XX			
XX	30-JUN-2000; 2000US-216503P.		
XX			
XX	(MILL-) MILLENNIUM PHARM INC.		
XX			
XX	Glucksmann MA;		
XX			
XX	WPI; 2002-140091/18.		
XX			
XX	P-PSDB; AAE16592.		
XX			
XX	New isolated human uridine kinase family polypeptide 57658, useful for		
XX	treating hematopoietic neoplastic disorders and disorders of neurons,		
XX	heart and blood vessels -		
XX			

PS Claim 1; Fig 1a; 103pp; English.

XX The patent discloses human uridine kinase-like polypeptides, designated
CC 57658 and polynucleotides encoding such proteins. 57658 DNAs and proteins
CC are useful for developing diagnostic and therapeutic agents for 57658-
CC mediated or related disorders such as haematopoietic neoplastic disorders
CC (e.g. leukaemia), haematopoietic disorders (e.g. psoriasis, dermatitis,
CC diabetes mellitus, rheumatoid arthritis, disorders of bone (e.g. rickets,
CC osteoporosis), heart (e.g. sarcoma, myocardial infarction), blood vessels
CC (e.g. hypertension, atherosclerosis) and neurons (e.g. Huntington's
CC disease, Alzheimer's disease, Parkinson's disease). 57658 sequences or
CC their antibodies are useful in screening assays, detection assays (e.g.
CC forensic biology) and predictive medicine (e.g. diagnostic assays,
CC prognostic assays, and monitoring clinical trials and pharmacogenomics).
CC They are useful as reagents for diagnosing and treating 57658-mediated
CC disorders. 57658 DNAs are useful in gene therapy, in chromosome mapping,
CC to identify an individual from a minute biological sample (tissue typing)
CC and to aid in forensic identification of the biological sample. The
CC present sequence is a DNA encoding human 57658 protein.

XX Sequence 1624 BP; 354 A; 427 C; 506 G; 337 T; 0 other;

Query Match 100.0%; Score 834; DB 24; Length 1624;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 834; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCTTCGGCGGGAGCGGAGACTGCGAGAGCCCGCGCGGAGGCCGACCTCCGAC 60
DB ATGGCTTCGGCGGGAGCGGAGACTGCGAGAGCCCGCGCGGAGGCCGACCTCCGAC 153
QY 61 CAGCGGCCCTTCTGATAGGGGTGAGCGCGGCACTCCAGCGGAAAGTGCACCTGTGT 120
DB CAGCGGCCCTTCTGATAGGGGTGAGCGCGGCACTCCAGCGGAAAGTGCACCTGTGT 213
QY 121 GAGAGATCATGAGTGTCTGGACAGAACGAGGTGAGCGCGGAGCGGAAAGTGTCT 180
DB GAGAGATCATGAGTGTCTGGACAGAACGAGGTGAGCGCGGAGCGGAAAGTGTCT 273
QY 181 ATCTGAGCGGAGCAGGTTCTCAAGGTCCTGACCGCAGAGCAAGAGCGGCTTG 240
DB ATCTGAGCGGAGCAGGTTCTCAAGGTCCTGACCGCAGAGCAAGAGCGGCTTG 333
QY 241 AAAGCAGATCAATTTTGCATCCAGATGCTTTGATATGATTTGATGACAGAGCT 300
DB AAAGCAGATCAATTTTGCATCCAGATGCTTTGATATGATTTGATGACAGAGCT 333
QY 301 CTGAGAGAACATCGTGAAGGCGAAACGCTGAGGTGCGCATATGATTTTGTGACAC 360
DB CTGAGAGAACATCGTGAAGGCGAAACGCTGAGGTGCGCATATGATTTTGTGACAC 453
QY 361 TCAAGTTTCCAGAGACCGGTGCTACCCCTGCGGACGTGTTCTGTTGAGGGCTC 420
DB TCAAGTTTCCAGAGACCGGTGCTACCCCTGCGGACGTGTTCTGTTGAGGGCTC 513
QY 421 TTGGTTTCTACAGCAGGAGATCCGAGCATGTTCCACTGCGGCTCTTGTGACAC 480
DB TTGGTTTCTACAGCAGGAGATCCGAGCATGTTCCACTGCGGCTCTTGTGACAC 573
QY 514 TTGGTTTCTACAGCAGGAGATCCGAGCATGTTCCACTGCGGCTCTTGTGACAC 573
DB TTGGTTTCTACAGCAGGAGATCCGAGCATGTTCCACTGCGGCTCTTGTGACAC 573
QY 481 GACTCCGAGCTCGAGGTCTCTCGAAGAGTTCTCCGAGACGTGCGCAGAGGAGGACCTG 540
DB GACTCCGAGCTCGAGGTCTCTCGAAGAGTTCTCCGAGACGTGCGCAGAGGAGGACCTG 633
QY 541 GAGCAATTCGAGCGAGTACACCACTTCTGTAACCCGCGCTTCAAGAGTTCTGCTG 600
DB GAGCAATTCGAGCGAGTACACCACTTCTGTAACCCGCGCTTCAAGAGTTCTGCTG 693
QY 601 CCGACAAGAGTATCCGATGATCATCCACGAGAGTGAACAATATGTTGTCATC 660
DB CCGACAAGAGTATCCGATGATCATCCACGAGAGTGAACAATATGTTGTCATC 753
QY 661 AACCTGATGTGACAGACATCCAGACATTTGTAATGTGACATCTGCAATGGACCGA 720
DB AACCTGATGTGACAGACATCCAGACATTTGTAATGTGACATCTGCAATGGACCGA 813

QY 721 GGAGGATCCAAATGGGGGAGCTTACAGCGGACCTTTCTGAGCCAGGGAGCACCTCGG 780
DB GGAGGATCCAAATGGGGGAGCTTACAGCGGACCTTTCTGAGCCAGGGAGCACCTCGG 873
QY 781 ATGCTGACCTTGGCAAGCGTACATTTTGAATCCAGACGAGACCCCACTGA 834
DB ATGCTGACCTTGGCAAGCGTACATTTTGAATCCAGACGAGACCCCACTGA 927

RESULT 3

AAH15621 standard; cDNA; 2160 BP.

XX AAH15621.
XX 26-JUN-2001 (first entry)

DE Human cDNA sequence SEQ ID NO:13951.

KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; se.

OS Homo sapiens.

PN EP1074617-A2.

PD 07-FEB-2001.

PF 28-JUL-2000; 2000BP-0116126.

PR 29-JUL-1999; 99JP-0248036.

PR 27-AUG-1999; 99JP-0300253.

PR 11-JUN-2000; 2000JP-0118776.

PR 02-MAY-2000; 2000JP-0183767.

PR 09-JUN-2000; 2000JP-0241899.

XX (HELI-) HELIX RES INSE.

XX Ota T, Isegai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX WPI; 2001-318749/34.

PT Primer sets for synthesizing polynucleotides, particularly the 5602

PT full-length cDNAs defined in the specification, and for the detection

PT and/or diagnosis of the abnormality of the proteins encoded by the

PT full-length cDNAs -

PS Claim 8; SEQ ID 13951; 2537pp + CD ROM; English.

XX The present invention describes primer sets for synthesizing 5602

CC full-length cDNAs defined in the specification. Where a primer set

CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary

CC to the complementary strand of a polynucleotide which comprises one of

CC the 5602 nucleotide sequences defined in the specification, where the

CC oligonucleotide comprises at least 15 nucleotides; and (b) a combination

CC of an oligonucleotide comprising a sequence complementary to the

CC complementary strand of a polynucleotide which comprises a 5'-end

CC sequence and an oligonucleotide comprising a sequence complementary to a

CC polynucleotide which comprises a 3'-end sequence, where the

CC oligonucleotide comprises at least 15 nucleotides and the combination of

CC the 5'-end sequence/3'-end sequence is selected from those defined in

CC the specification. The primer sets can be used in antisense therapy and

CC in gene therapy. The primers are useful for synthesizing polynucleotides,

CC particularly full-length cDNAs. The primers are also useful for the

CC detection and/or diagnosis of the abnormality of the proteins encoded by

CC the full-length cDNAs. The primers allow obtaining of the full-length

CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and

CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to

CC AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632

CC represent oligonucleotides, all of which are used in the exemplification

CC of the present invention.

Sequence 2160 BP; 457 A; 591 C; 671 G; 441 T; 0 other;

Query Match 100.0%; Score 834; DB 22; Length 2160;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 834; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 ATGGCTTCGGCGGAGCGAAGACTGAGAGAGCCCGCGCGAGCGCCGACCCGCGCAC 60
   |||
DB 95 ATGGCTTCGGCGGAGCGAAGACTGAGAGAGCCCGCGCGAGCGCCGACCCGCGCAC 154
QY 61 CAGCGGCCCTTCCTGATAGGGGTGAGCGCGGCACTCCAGCGGGAAGTGCACCGTGTGT 120
   |||
DB 155 CAGCGGCCCTTCCTGATAGGGGTGAGCGCGGCACTCCAGCGGGAAGTGCACCGTGTGT 214
QY 121 GAGAAATCATGAGTTCTTGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
   |||
DB 215 GAGAAATCATGAGTTCTTGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 274
QY 181 ATCCGAGCAGAGCAGAGTTCTCAAGAGTCTGACCGGAGAGAGAGAGAGAGAGAG 240
   |||
DB 275 ATCCGAGCAGAGCAGAGTTCTCAAGAGTCTGACCGGAGAGAGAGAGAGAGAGAG 334
QY 241 AAAGACAGTACATTTTGAACATCCAGATGCTTTGATATGATTTGATGACAGAGACT 300
   |||
DB 335 AAAGACAGTACATTTTGAACATCCAGATGCTTTGATATGATTTGATGACAGAGACT 394
QY 301 CTGAAGAACATCTGAGAGGCAAAAGCGTGAAGGTCCGACCTATGATTTTGTGACACAC 360
   |||
DB 395 CTGAAGAACATCTGAGAGGCAAAAGCGTGAAGGTCCGACCTATGATTTTGTGACACAC 454
QY 361 TCAGAGTTACAGAGACAGCGTGTCTACCCCTGCGAGCGGTCTGTTGAGGGGATC 420
   |||
DB 455 TCAGAGTTACAGAGACAGCGTGTCTACCCCTGCGAGCGGTCTGTTGAGGGGATC 514
QY 421 TTGGTGTCTTACAGCAGAGATCCGGGACATGTTTCCACCTGCGCTCTTCTGTGACACC 480
   |||
DB 515 TTGGTGTCTTACAGCAGAGATCCGGGACATGTTTCCACCTGCGCTCTTCTGTGACACC 574
QY 481 GACTCCGACGTGAGCTGTCTCGAAGAGTTCTCCGGGACGTGCGCGAGAGGAGGACCTG 540
   |||
DB 575 GACTCCGACGTGAGCTGTCTCGAAGAGTTCTCCGGGACGTGCGCGAGAGGAGGACCTG 634
QY 541 GAGAGATTCTGAGCAGATCACCACTTGTGTAAGCGGCTTTCGAGGAGTTCTGCTG 600
   |||
DB 635 GAGAGATTCTGAGCAGATCACCACTTGTGTAAGCGGCTTTCGAGGAGTTCTGCTG 694
QY 601 CCGCAAAAGAGTATGCGCATGTGATCATCCACAGAGAGTGAACAATATGTTGCCATC 660
   |||
DB 695 CCGCAAAAGAGTATGCGCATGTGATCATCCACAGAGAGTGAACAATATGTTGCCATC 754
QY 661 AACCTGATCGTGCAGACATCCAGAGCAATTTGTAATGATCATCTGCAATATGCGACCGA 720
   |||
DB 755 AACCTGATCGTGCAGACATCCAGAGCAATTTGTAATGATCATCTGCAATATGCGACCGA 814
QY 721 GAGAGGTCATGAGGCGGAGGAGTCAAGCGGACCTTTTCTGAGCGAGGGGACACCCCTGGG 780
   |||
DB 815 GAGAGGTCATGAGGCGGAGGAGTCAAGCGGACCTTTTCTGAGCGAGGGGACACCCCTGGG 874
QY 781 ATGTGACCTCTGCAAAAGGTGCAATTTGAGAGTCCAGAGCAGAGACCCCACTGA 834
   |||
DB 875 ATGTGACCTCTGCAAAAGGTGCAATTTGAGAGTCCAGAGCAGAGACCCCACTGA 928

```

RESULT 4
ID AAK8735
AAK8735 standard; DNA; 834 BP.

AC AAK98735;
DT 02-MAY-2002 (first entry)
XX DNA of a human uridine kinase (UDK).
DE
XX

KW Human; uridine kinase; diagnostic assay; mutation detection; UDK;
KW probe; chromosome localization study; tissue expression; gene therapy;
KW antibody; vaccine; human ovarian cancer; immunological disorder;
KW human colon carcinoma; immunogen; ds.
OS Homo sapiens.

Key Location/Qualifiers
FT CDS 1..780
FT /*tag= a
FT /partial
FT /note= "No stop codon"

MO200172863-A2.

04-OCT-2001.

27-MAR-2001; 2001MO-US09663.

27-MAR-2000; 2000US-0536647.

(SMIK) SMITHKLINE BEECHAM CORP.

Ho YS, Johnson RK;

WPI; 2001-626259/72.

P-PSDB; AAO14412.

Novel human uridine kinase polypeptides useful for treating cancers,
PT and to identify agonists and antagonists of the polypeptide useful for
PT treating conditions associated with uridine kinase imbalance -
PS Claim 5; Page 22-23; 31pp; English.

XX The invention relates to newly identified human uridine kinase (UDK)
CC polypeptides and polynucleotides and methods for producing such
CC polypeptides by recombinant techniques. Also disclosed in the invention
CC are methods for utilizing uridine kinase polypeptides and polynucleotides
CC in diagnostic assays. The polynucleotides and polypeptides of the
CC invention may be used as diagnostic reagents by detecting mutations in an
CC associated gene. An array of oligonucleotide probes comprising the
CC uridine kinase polynucleotide sequence or fragments thereof can be
CC constructed to conduct efficient screening of genetic mutations, for
CC example. Detection of abnormally decreased or increased levels of
CC polypeptide or mRNA expression may also be used for diagnosing or
CC determining susceptibility of a subject to a disease of the invention.
CC The polynucleotide sequences of the invention can be used for chromosome
CC localization studies and tissue expression studies. The polypeptides of
CC the invention or fragments thereof may be used as immunogens to produce
CC antibodies. These antibodies may be employed to isolate or identify
CC clones expressing the polypeptide. The polypeptides and polynucleotides
CC of the invention can be used as a vaccine or in gene therapy to treat
CC diseases such as human ovarian cancer, human colon carcinomas, and
CC immunological disorders. This polynucleotide sequence represents the DNA
CC of a human uridine kinase of the invention.

Sequence 834 BP; 194 A; 219 C; 265 G; 156 T; 0 other;

Query Match 93.9%; Score 783; DB 23; Length 834;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 833; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 ATGGCTTCGGCGGAGCGAAGACTGAGAGAGCCCGCGCGAGCGCCGACCCGCGCAC 60
   |||
DB 1 ATGGCTTCGGCGGAGCGAAGACTGAGAGAGCCCGCGCGAGCGCCGACCCGCGCAC 60
QY 61 CAGCGGCCCTTCCTGATAGGGGTGAGCGCGGCACTCCAGCGGGAAGTGCACCGTGTGT 120
   |||
DB 61 CAGCGGCCCTTCCTGATAGGGGTGAGCGCGGCACTCCAGCGGGAAGTGCACCGTGTGT 120
QY 121 GAGAAATCATGAGTTCTTGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
   |||
DB 121 GAGAAATCATGAGTTCTTGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180

```


QY 181 ATCTGAGCCAGGAGGTTCTCAAGAGTCTGACGCGCAGAGGAGGCGCAAGGCGCTTG 240
 DB 181 ATCTGAGCCAGGAGGAGGTTCTCAAGAGTCTGACGCGCAGAGGAGGCGCAAGGCGCTTG 240
 QY 241 AAGAGCAGTACATTTTGAACATCCAGATCCCTTTGATATGATTTGAATCAAGAGACT 300
 DB 241 AAGAGCAGTACATTTTGAACATCCAGATCCCTTTGATATGATTTGAATCAAGAGACT 300
 QY 301 CTGAAGAACATCTGAGGGGCAAAACGCTGAGAGTCCGACCTATGATTTTGTGACACAC 360
 DB 301 CTGAAGAACATCTGAGGGGCAAAACGCTGAGAGTCCGACCTATGATTTTGTGACACAC 360
 QY 361 TCAAGGTTACAGAGCAGAGATCCGAGATGTTCCACCTGCGCTCTTCTGTGACACC 420
 DB 361 TCAAGGTTACAGAGCAGAGATCCGAGATGTTCCACCTGCGCTCTTCTGTGACACC 420
 QY 421 TTGGTGTCTACAGCCAGAGATCCGAGATGTTCCACCTGCGCTCTTCTGTGACACC 480
 DB 421 TTGGTGTCTACAGCCAGAGATCCGAGATGTTCCACCTGCGCTCTTCTGTGACACC 480
 QY 481 GACTCCGAGCTCAGGCTGTCTCGAAGATTCTCCGGAGCTGGCGGAGGAGGAGCTTG 540
 DB 481 GACTCCGAGCTCAGGCTGTCTCGAAGATTCTCCGGAGCTGGCGGAGGAGGAGCTTG 540
 QY 541 GAGCAGATTCTGACGAGTACACCACTTGTAAGCGGCGCTTCCAGAGAGTTCTGCTG 600
 DB 541 GAGCAGATTCTGACGAGTACACCACTTGTAAGCGGCGCTTCCAGAGAGTTCTGCTG 600
 QY 601 CCGACAAAGATATGCGGATGATGATCCACGAGAGTGGACATATGTTGCCATC 660
 DB 601 CCGACAAAGATATGCGGATGATGATCCACGAGAGTGGACATATGTTGCCATC 660
 QY 661 AACCTGATCTGACGACATCCAGAGATCTGTAATGTAATCTGCAATGCGACCGA 720
 DB 661 AACCTGATCTGACGACATCCAGAGATCTGTAATGTAATCTGCAATGCGACCGA 720
 QY 721 GAGAGGTCATGAGGCGGAGCTCAAGCGGAGCCCTTTCTAGGCGAGGCGACCCCTGG 780
 DB 721 GAGAGGTCATGAGGCGGAGCTCAAGCGGAGCCCTTTCTAGGCGAGGCGACCCCTGG 780
 QY 781 ATCTGACCTCTGCGCAACCGGTCAATTTGAGTCCAGAGACCCCACTGA 834
 DB 781 ATCTGACCTCTGCGCAACCGGTCAATTTGAGTCCAGAGACCCCACTGA 834
 RESULT 5
 AA160444
 ID AA160444 standard; cDNA; 2152 BP.
 AC AA160444;
 XX
 DT 22-OCT-2001 (first entry)
 DE Human polynucleotide SEQ ID NO 4433.
 XX
 XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KM peripheral nervous system; neuropathy; central nervous system; CNS;
 KM Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KM amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KM chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 KM leukemia; ss.
 XX
 OS Homo sapiens.
 PN WO200153312-A1.
 XX
 PD 26-JUL-2001.
 XX
 PF 26-DEC-2000; 2000MO-US34263.
 XX
 PR 21-JAN-2000; 2000US-048725.
 PR 25-APR-2000; 2000US-0552317.

PR 09-JUL-2000; 2000US-0598042.
 PR 19-JUL-2000; 2000US-0620312.
 PR 03-AUG-2000; 2000US-0653450.
 PR 14-SEP-2000; 2000US-0662191.
 PR 19-OCT-2000; 2000US-0693036.
 PR 29-NOV-2000; 2000US-0727344.
 XX
 PA (HYSE-) HYSO INC.
 XX
 PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Weinman T, Xu C, Xue AJ, Yang Y, Zhang J;
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
 XX
 DR WPI: 2001-442253/47.
 DR P-PSDB; AAM41288.
 XX
 PT Novel nucleic acids and polypeptides, useful for treating disorders
 PT such as central nervous system injuries -
 PS
 PS Claim 1; SEQ ID NO 4433; 10078bp; English.
 XX
 CC The invention relates to human nucleic acids (AA157798-AA161369) and
 CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
 CC immunosuppressant and cyostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemias and
 CC C.N.S disorders.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification.
 XX
 SQ Sequence 2152 BP; 468 A; 581 C; 660 G; 442 T; 1 other;
 Query Match 93.9%; Score 783; DB 22; Length 2152;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 833; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 ATGCTCTGCGCGGAGCGCAAGACTGCGAGAGCCCGCGGAGCGCAAGCTCCGCAC 60
 DB 59 ATGGCTTCGCGGAGGCGAAGACTGCGAGAGCCCGCGGAGCGCAAGCTCCGCAC 118
 QY 61 CAGCGGCCCTTCTGATGAGGAGTGAAGCGGCGCACTGCGAGGAGGAGTCAAGCGGTGT 120
 DB 119 CAGCGGCCCTTCTGATGAGGAGTGAAGCGGCGCACTGCGAGGAGGAGTCAAGCGGTGT 178
 QY 121 GAGAGATCATGAGAGTGTGCGGAGCAAGAGGTGAACAGCGGAGCGGAAGGTGTC 180
 DB 179 GAGAGATCATGAGAGTGTGCGGAGCAAGAGGTGAACAGCGGAGCGGAAGGTGTC 238
 QY 181 ATCTGAGCCAGAGCAGGTTCTTCAAGAGTCTGACGCGCAGAGCAAGAGGCGCTTG 240
 DB 239 ATCTGAGCCAGAGCAGGTTCTTCAAGAGTCTGACGCGCAGAGCAAGAGGCGCTTG 298
 QY 241 AAGAGCAGTACATTTTGAACATCCAGATCCCTTTGATATGATTTGAATGACAGAGACT 300
 DB 299 AAGAGCAGTACATTTTGAACATCCAGATCCCTTTGATATGATTTGAATGACAGAGACT 358
 QY 301 CTGAAGAACATCTGAGGGGCAAAACGCTGAGAGTCCGACCTATGATTTTGTGACACAC 360
 DB 359 CTGAAGAACATCTGAGGGGCAAAACGCTGAGAGTCCGACCTATGATTTTGTGACACAC 418
 QY 361 TCAAGGTTACAGAGCAGGAGTGTCTTACCTGCGAGAGTGTGTTTGAAGGCGATC 420
 DB 419 TCAAGGTTACAGAGCAGGAGTGTCTTACCTGCGAGAGTGTGTTTGAAGGCGATC 478
 QY 421 TTGGTGTCTACAGCCAGAGATCCGAGATGTTCCACCTGCGCTCTTCTGTGACACC 480

```

Db      479  TTGGGTTCTACAGCCAGAGATCCGGGACATGTTCCACCTGCGCTCTTGTGGACACC 538
Qy      481  GACTCCGACCTCAGGCTCTTCGAAGATTTCTCCGGACCTGCCCGGAGGAGGACCTG 540
Db      539  GACTCCGACCTCAGGCTCTTCGAAGATTTCTCCGGACCTGCCCGGAGGAGGACCTG 598
Qy      541  GAGCAGATTCTGACGACATACCACTTCGGAAGCCGGCTTCGAGAGATTCTGACCTG 600
Db      599  GAGCAGATTCTGACGACATACCACTTCGGAAGCCGGCTTCGAGAGATTCTGACCTG 658
Qy      601  CCGAAGAAGATATGCGATGATGATGATGATGATGATGATGATGATGATGATGATG 660
Db      659  CCGAAGAAGATATGCGATGATGATGATGATGATGATGATGATGATGATGATGATG 718
Qy      661  AACTGATTCGTCGACGACATCCAGACATTTCTGAATGATGATGATGATGATGATG 720
Db      719  AACTGATTCGTCGACGACATCCAGACATTTCTGAATGATGATGATGATGATGATG 778
Qy      721  GGAGGGTTCATGAGCGGAGCTACCAAGCGGACCTTTCTGAGCGGAGGAGGACCTG 780
Db      779  GGAGGGTTCATGAGCGGAGCTACCAAGCGGACCTTTCTGAGCGGAGGAGGACCTG 838
Qy      781  ATGCTGACCTCTGCGAAACGGTCACTTTGAGTTCAGCAGCAGACCCCACTGA 834
Db      839  ATGCTGACCTCTGCGAAACGGTCACTTTGAGTTCAGCAGCAGACCCCACTGA 892

```

RESULT 6
AAH04832
ID AAH04832 standard; cDNA; 753 BP.

AAH04832;

26-JUN-2001 (first entry)

Human cDNA clone (5'-primer) SEQ ID NO:1667.

Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

Homo sapiens.

EP1074617-A2.

07-FEB-2001.

28-JUL-2000; 2000EP-0116126.

29-JUL-1999; 99JP-0248036.

27-AUG-1999; 99JP-0300253.

11-JAN-2000; 2000JP-0118776.

02-MAY-2000; 2000JP-0183767.

09-JUN-2000; 2000JP-0241899.

(HELI-) HELIX RES INST.

Oca T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

WPI; 2001-318749/34.

Primer sets for synthesizing polynucleotides, particularly the 5602

full-length cDNAs defined in the specification, and for the detection

and/or diagnosis of the abnormality of the proteins encoded by the

full-length cDNAs -

Claim 1; SEQ ID 1667; 2537bp + CD ROM; English.

The present invention describes primer sets for synthesizing 5602

full-length cDNAs defined in the specification. Where a primer set

comprises: (a) an oligo-dT primer and an oligonucleotide complementary

to the complementary strand of a polynucleotide which comprises one of

the 5602 nucleotide sequences defined in the specification, where the

oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesizing polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention.

Sequence 753 BP; 158 A; 196 C; 261 G; 135 T; 3 other;

Query Match 68.9%; Score 575; DB 22; Length 753;

Best Local Similarity 100.0%; Pred.No.1.6e-289; Mismatches 0; Gaps 0;

Matches 575; Conservative 0; Indels 0; Gaps 0;

```

Qy      1  ATGGCTTCGCGGAGGCGAAGACTGCGAGAGCCCGCGCGGAGGCGGACCTCCGAC 60
Db      95  ATGGCTTCGCGGAGGCGAAGACTGCGAGAGCCCGCGCGGAGGCGGACCTCCGAC 154
Qy      61  CAGCGGCCCTTCTGATGAGGGGTGAGCGGCGGCACTGCGGCGGAAAGTGCACCTGTGT 120
Db      155  CAGCGGCCCTTCTGATGAGGGGTGAGCGGCGGCACTGCGGCGGAAAGTGCACCTGTGT 214
Qy      121  GAGAAGATCAATGAGTTCGCGGAGCAAGACGAGTGGAAAGGCGGAGGAGTGTGTC 180
Db      215  GAGAAGATCAATGAGTTCGCGGAGCAAGACGAGTGGAAAGGCGGAGGAGTGTGTC 274
Qy      181  ATCCTGAGCCAGAGCAGGTTCTCAAGGTCCTGACGCGCAGAGCAGAAAGGCCAGCTTG 240
Db      275  ATCCTGAGCCAGAGCAGGTTCTCAAGGTCCTGACGCGCAGAGCAGAAAGGCCAGCTTG 334
Qy      241  AAAAGCAGTACAAATTTTGACATCCAGATGCTTTGATATGATTTGATGACAGGACT 300
Db      335  AAAAGCAGTACAAATTTTGACATCCAGATGCTTTGATATGATTTGATGACAGGACT 394
Qy      301  CTGAAGAACAATCGTGAAGGCGAAAGCGTGAAGTCCGACCTATGATTTTGTGACAC 360
Db      395  CTGAAGAACAATCGTGAAGGCGAAAGCGTGAAGTCCGACCTATGATTTTGTGACAC 454
Qy      361  TCAAGTTACAGAGACACAGGTGTCTACCTCGCGAGCGTGTCTGTTGAGGCGATC 420
Db      455  TCAAGTTACAGAGACACAGGTGTCTACCTCGCGAGCGTGTCTGTTGAGGCGATC 514
Qy      421  TTGGTGTTCACACGCAAGATCGGAGATTTCCACTGCGCTCTTCTGTGACAC 480
Db      515  TTGGTGTTCACACGCAAGATCGGAGATTTCCACTGCGCTCTTCTGTGACAC 574
Qy      481  GACTCGAGGTGAGGCTGCTCGAAGATTCCTCGGAGAGTGTGCGGAGGAGGAGGAGCTG 540
Db      575  GACTCGAGGTGAGGCTGCTCGAAGATTCCTCGGAGAGTGTGCGGAGGAGGAGGAGCTG 634
Qy      541  GAGCAGATTCTGACGACATACCACTTCGGAAGCCGGCTTCGAGAGATTCTGACCTG 575
Db      635  GAGCAGATTCTGACGACATACCACTTCGGAAGCCGGCTTCGAGAGATTCTGACCTG 669

```

RESULT 7
AAH75355
ID AAH75355 standard; cDNA; 900 BP.

AAH75355;

02-OCT-2001 (first entry)

DE Human uridine kinase encoding cDNA.
 XX Human; uridine kinase; UK; ss.
 XX Homo sapiens.
 XX Key Location/Qualifiers
 FT CDS 8..838
 FT /product= "uridine kinase"
 FT /note= "claimed in claim 1"
 FT
 XX CN1287172-A.
 XX 14-MAR-2001.
 XX PD
 XX PF 07-SEP-1999; 99CN-0118818.
 XX PR 07-SEP-1999; 99CN-0118818.
 XX PA (UYFU-) UNIV FUDAN.
 XX PI Yu L, Zhao Y, Zhang H;
 XX WPI; 2001-409529/44.
 XX P-PSDB; AAG64506.
 XX DR
 XX PT Human uridine kinase and its coding sequence, preparation and application -
 XX Claim 1; Page 15(Disclosure); 20pp; Chinese.
 XX The invention relates to human uridine kinase (UK).
 CC Sequence 900 BP; 201 A; 237 C; 292 G; 170 T; 0 other;
 SQ
 Query Match 62.0%; Score 517; DB 22; Length 900;
 Best Local Similarity 99.8%; Pred. No. 3e-259;
 Matches 567; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 172 AAGGTGTCATCTCTGAGCCAGAGAGGTTCTACAGGTCTCTGACGAGAGAGAGAGAGCC 231
 DB 176 AAGGTGTCATCTCTGAGCCAGAGAGGTTCTACAGGTCTCTGACGAGAGAGAGAGAGCC 235
 QY 232 AAGGCTTGAAGAGACGTACATTTTGACCATTCAGATGCTTTGATATGATTTGATG 291
 DB 236 AAGGCTTGAAGAGACGTACATTTTGACCATTCAGATGCTTTGATATGATTTGATG 295
 QY 292 CACAGACTCTGAAGAATCGTGGAGGCAAAACGGTGGAGGTGCCAGCTATGATTTT 351
 DB 296 CACAGACTCTGAAGAATCGTGGAGGCAAAACGGTGGAGGTGCCAGCTATGATTTT 355
 QY 352 GTGACACTCAAGTTACAGAGACCAAGTGTCTACCTCGGAGCGTGTCTGTTT 411
 DB 356 GTGACACTCAAGTTACAGAGACCAAGTGTCTACCTCGGAGCGTGTCTGTTT 415
 QY 412 GAGGAGCTTTGTGTCTACAGCCAGAGATCCGGACATGTTCCACTGCGCTTTC 471
 DB 416 GAGGAGCTTTGTGTCTACAGCCAGAGATCCGGACATGTTCCACTGCGCTTTC 475
 QY 472 GTGAGACCACTCCGAGTCCAGGCTGTCTGAAGAGTTCTCCGGAGCGTGGCCGAGG 531
 DB 476 GTGAGACCACTCCGAGTCCAGGCTGTCTGAAGAGTTCTCCGGAGCGTGGCCGAGG 535
 QY 532 AGGAGCTTGAGAGAGATTTGACGACAGTACCACTTCTGTAAGCCGCGCTTGAAGAG 591
 DB 536 AGGAGCTTGAGAGAGATTTGACGACAGTACCACTTCTGTAAGCCGCGCTTGAAGAG 595
 QY 592 TTTCGCTGCGGCAAGAGATGCGAGTGTATCCCAAGAGAGTGGACATATG 651
 DB 596 TTTCGCTGCGGCAAGAGATGCGAGTGTATCCCAAGAGAGTGGACATATG 655
 QY 652 GTTGCCATCAACTGATGTGACACATCAGGACATCTGATGATGATCTGCAAA 711

DB 656 GTTGCCATCAACTGATGTGACACATCAGGACATTTCTGATGTGACATCTGCAAA 715
 QY 712 TGGCACCAGAGAGGTTCATGAGCGGA 739
 DB 716 TGGCACCAGAGAGGTTCATGAGCGGA 743
 RESULT 8
 ID ABL89762 standard; cDNA; 1396 BP.
 AC ABL89762;
 XX 24-MAY-2002 (first entry)
 DE Human polynucleotide SEQ ID NO 324.
 XX Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
 KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor;
 KW vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
 KW cardiact; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; human; secreted protein; gene; ss.
 XX Homo sapiens.
 OS
 XX WO200190304-A2.
 XX 29-NOV-2001.
 XX PD
 XX PF 18-MAY-2001; 2001WO-US16450.
 XX PR 19-MAY-2000; 2000US-205515P.
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA Birse CE, Rosen CA;
 PI WPI; 2002-122018/16.
 XX P-PSDB; ABB89353.
 DR
 XX Novel 1405 isolated polypeptides, useful for diagnosis, treatment and prevention of neural, immune system, muscular, reproductive, gastrointestinal, pulmonary, cardiovascular, renal and proliferative disorders -
 PT Claim 4; SEQ ID NO 324; 2081bp + Sequence Listing; English.
 PS
 XX The invention relates to novel genes (AB189449-AB190853) and proteins (AB189449-AB190444) useful for preventing, treating or ameliorating isolated conditions e.g. by protein or gene therapy. The genes are isolated from a range of human tissues disclosed in the specification. The nucleic acids, proteins, antibodies and (ant)agonists are useful in the diagnosis, treatment and prevention of: (a) cancer; e.g. breast and ovarian cancer and other cancers of the adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune disorders e.g. Addison's disease, allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c) cardiovascular disorders such as myocardial ischaemias; (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f) infectious diseases such as viral, bacterial, fungal and parasitic infections.
 CC Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 CC
 XX Sequence 1396 BP; 317 A; 358 C; 425 G; 287 T; 9 other;
 SQ
 Query Match 62.0%; Score 517; DB 24; Length 1396;
 Best Local Similarity 99.6%; Pred. No. 3e-259;
 Matches 737; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

```

QY 1 ATGGCTTCGGCGGAGGAGAAAGTGCAGAGAGCCCGCGCGAGAGCCGACCGTCCGAC 60
XX |||||||
Db 41 ATGGCTTCGGCGGAGGAGAAAGTGCAGAGAGCCCGCGCGAGAGCCGACCGTCCGAC 100
QY 61 CAGCGGCCCTTCTTGATAGGGGTGAGCGCGGCACTGCGCAGCGGGAAGTGCACCGTGTGT 120
XX |||||||
Db 101 CAGCGGCCCTTCTTGATAGGGGTGAGCGCGGCACTGCGCAGCGGGAAGTGCACCGTGTGT 160
QY 121 GAGAAATCATGAGATTTCTGGGACAGAAACGAGGTGGAACCGCGCACCGGAAAGTGTCT 180
XX |||||||
Db 161 GAGAAATCATGAGATTTCTGGGACAGAAACGAGGTGGAACCGCGCACCGGAAAGTGTCT 220
QY 181 ATCTGAGCCAGAGCAGGTTCTACAAAGTCTCTGACGAGAGAGAGCCAAAGCCCTTG 240
XX |||||||
Db 221 ATCTGAGCCAGAGCAGGTTCTACAAAGTCTCTGACGAGAGAGAGCCAAAGCCCTTG 280
QY 241 AAAGACAGTACAAATTTTGACCATCCAGATGCTTTGATTAATGATTTGATGCACAGACT 300
XX |||||||
Db 281 AAAGACAGTACAAATTTTGACCATCCAGATGCTTTGATTAATGATTTGATGCACAGACT 340
QY 301 CTGAAGAACATCTGAGAGGCGAAACCGGTGAGAGTCCGACCTATGATTTTGACACAC 360
XX |||||||
Db 341 CTGAAGAACATCTGAGAGGCGAAACCGGTGAGAGTCCGACCTATGATTTTGACACAC 400
QY 361 TCAAGTTACAGAGACACCGGTGCTACCTCGGACGATGTTCTGTTGAGGCGATC 420
XX |||||||
Db 401 TCAAGTTACAGAGACACCGGTGCTACCTCGGACGATGTTCTGTTGAGGCGATC 460
QY 421 TTGGTGTCTACAGCAGAGAGATCCGGGACATGTTCCACTGCGCTCTTCTGAGACAC 480
XX |||||||
Db 461 TTGGTGTCTACAGCAGAGAGATCCGGGACATGTTCCACTGCGCTCTTCTGAGACAC 520
QY 481 GACTCCGACGTCAGGCTGTCTGAAAGATTTCCGGGACGTCGCGCGAGGAGGAGACCTG 540
XX |||||||
Db 521 GACTCCGACGTCAGGCTGTCTGAAAGATTTCCGGGACGTCGCGCGAGGAGGAGACCTG 580
QY 541 GAGCAGATTTGACGACGATACACCACTTCTGTAAGCCGCGCTTCAAGAGATTTGACCTG 600
XX |||||||
Db 581 GAGCAGATTTGACGACGATACACCACTTCTGTAAGCCGCGCTTCAAGAGATTTGACCTG 639
QY 601 CCGACAAAGAGTATGCCGATGTGATCATCCACGAGAGTGAACAATATGTTGCCATC 660
XX |||||||
Db 640 CCGACAAAGAGTATGCCGATGTGATCATCCACGAGAGTGAACAATATGTTGCCATC 699
QY 661 AACCTGATCGTGCAGACATCCAGAGCACTTCTGTAATGTTGACATCTGAAATGACACCGA 720
XX |||||||
Db 700 AACCTGATCGTGCAGACATCCAGAGCACTTCTGTAATGTTGACATCTGCAAAATGACACCGA 759
QY 721 GGAGGTTCCAAATGGCGCGAG 740
XX |||||||
Db 760 GGAGGTTCCAAATGGCGCGAG 779

RESULT 9
ABBS1801
ID ABBS1801 standard; cDNA; 734 BP.
XX
AC ABBS1801;
XX
DT 21-OCT-2002 (first entry)
XX
DE Human mdt cDNA Incyte ID No: LI:235557.12:2001JUN12.
XX
KW Human; molecule for disease detection and treatment; MDDR; cancer;
KW cell proliferative disorder; arteriosclerosis; cirrhosis; asthma;
KW autoimmune disorder; inflammatory disorder; Crohn's disease;
KW multiple sclerosis; cytostatic; antiarteriosclerotic; antiinflammatory;
KW hepatotropic; immunosuppressive; antiaschematic; gene therapy; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200255738-A2.
XX

```

```

PD 18-JUL-2002.
XX
PF 09-JAN-2002; 2002MO-US01008.
XX
PR 12-JAN-2001; 2001US-261622P.
XX
PR 16-JAN-2001; 2001US-261865P.
PR 17-JAN-2001; 2001US-262208P.
PR 17-JAN-2001; 2001US-262209P.
PR 17-JAN-2001; 2001US-262326P.
PR 19-JAN-2001; 2001US-263063P.
PR 19-JAN-2001; 2001US-263065P.
PR 19-JAN-2001; 2001US-263329P.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Panzer SR, Lincoln SE, Altus CM, Dufour GE, Hillman JL, Jones AL,
PI Dam TC, Liu TF, Harris B, Flores V, Dafo A, Marwaha R, Chen MJ,
PI Chang SC, Gerstein EH, Peralta CH, David MH, Lewis SA;
XX
DR WPI: 2002-590679/63.
DR P-PSDB: ABG70328.
XX
PT New disease detection and treatment molecule (MDPT) polynucleotides and
PT polypeptides, useful in diagnosing, studying, preventing or treating
PT diseases associated with MDDR expression, e.g. autoimmune or
PT inflammatory disorders -
XX
PS Claim 1; Page 98; 129pp; English.
XX
XX The present invention relates to the isolation of novel human
XX molecules for disease detection and treatment (MDPT), and the
XX polynucleotide sequences (mdpt) encoding them. The MDPT polypeptides
XX may be used to screen for molecules that bind to, or are bound by the
XX encoded polypeptides, and to develop a transcript image of a tissue or
XX cell type. Probes comprising at least 20 nucleotides of the mdpt
XX polynucleotide may be used to assess the toxicity of a test compound.
XX The MDPT polypeptides and mdpt polynucleotides are useful in the
XX diagnosis, study, prevention and treatment of diseases associated with
XX the expression of molecules for disease detection and treatment. Such
XX disorders include cell proliferative disorders (e.g. arteriosclerosis,
XX cirrhosis, or cancers), and autoimmune/inflammatory disorders
XX (e.g. asthma, Crohn's disease, or multiple sclerosis). The mdpt
XX polynucleotides may also be used as molecule markers, in microarrays,
XX and in somatic or germ-line gene therapy. ABBS1779-ABBS1814 encode
XX the MDPT proteins of the invention.
XX
SQ Sequence 734 BP; 166 A; 181 C; 226 G; 154 T; 7 other;
XX
Query Match 40.5%; Score 338; DB 24; Length 734;
Best Local Similarity 100.0%; Pred. No. 7.3e-166;
Matches 338; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 266 CAGATGCTTTGATTAATGATTTGATGCACAGACTCTGAAAGACATCTGAGAGGCAAAA 325
XX |||||||
Db 12 CAGATGCTTTGATTAATGATTTGATGCACAGACTCTGAAAGACATCTGAGAGGCAAAA 71
QY 326 CGGTGAGAGTCCGACCTATGATTTTGTGACACACTCAAGTTACAGAGACCAACCGTGG 385
XX |||||||
Db 72 CGGTGAGAGTCCGACCTATGATTTTGTGACACACTCAAGTTACAGAGACCAACCGTGG 131
QY 386 TCTACCTGCGAGAGTGTCTGTGTTGAGGAGCATCTTGGTGTCTACAGCAGAGATCC 445
XX |||||||
Db 132 TCTACCTGCGAGAGTGTCTGTGTTGAGGAGCATCTTGGTGTCTACAGCAGAGATCC 191
QY 446 GGGACATGTTTCCACTGCGCTCTTGTGAGACACCGACTCCGACGTCAAGCTGTCTGAA 505
XX |||||||
Db 192 GGGACATGTTTCCACTGCGCTCTTGTGAGACACCGACTCCGACGTCAAGCTGTCTGAA 251
QY 506 GAGTTCCTCGGAGCGTGGCCGAGGAGGAGCACTGAGACAGATTCTGACGCAATACCA 565
XX |||||||
Db 252 GAGTTCCTCGGAGCGTGGCCGAGGAGGAGCACTGAGACAGATTCTGACGCAATACCA 311
QY 566 CTTCTGTAAGCCGCGCTTCAAGAGTGTCTGCTTCCG 603
XX

```

Db 312 CCTTCGTGAAGCCGCGCTTCGAGGAGTTCGTGCGCCG 349
|||||
RESULT 10
AAL36439
ID AAL36439 standard; DNA; 9732 BP.
XX
AC AAL36439;
XX
DT 08-JAN-2002 (first entry)
DE Human musculoskeletal system related polynucleotide SEQ ID NO 2804.
XX
KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor;
KW vulnereary; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW neurological disease; infection; human; secreted protein;
KW musculoskeletal system; ds.
XX
OS Homo sapiens.
XX
PN W020015367-A1.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001W0-US01338.
XX
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.

PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235634.
PR 27-SEP-2000; 2000US-0235636.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.

PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251088.
PR 05-DEC-2000; 2000US-0251988.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251858.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
XX (HUMA-) HUMAN GENOME SCL INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-451937/48.
XX
XX Isolated polypeptide for treating, preventing and/or prognosing
PT disorders related to the musculoskeletal system including
PT musculoskeletal cancers and also for testing and detection e.g.
PT diagnosis -
XX
XX Example 2; SEQ ID NO 2804; 781bp + Sequence Listing; English.
XX
XX The invention relates to novel genes (AAL34669-AAL37666) and proteins
CC (AB03087-AB04109) associated with the musculoskeletal system useful
CC for preventing, treating or ameliorating medical conditions e.g. by
CC protein or gene therapy. The genes are isolated from a range of human
CC tissues disclosed in the specification. The nucleic acids, proteins,
CC antibodies and (ant)agonists are useful in the diagnosis, treatment,
CC and prevention of: (a) cancer, e.g. breast and ovarian cancer and
CC other cancers of the adrenal gland, bone, bone marrow, breast,
CC gastrointestinal tract, liver, lung, or urogenital; (b) immune
CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis;
CC (c) cardiovascular disorders such as myocardial ischaemia; (d) wound
CC healing; (e) neurological diseases such as cerebral anoxia and epilepsy;
CC and (f) infectious diseases such as viral, bacterial, fungal and
CC parasitic infections.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pat_sequences.
XX
XX Sequence 9732 BP; 2063 A; 2805 C; 2561 G; 2303 T; 0 other;
SQ
Query Match 21.9%; Score 183; DB 22; Length 9732;
Best Local Similarity 100.0%; Pred. No. 5.8e-85;
Matches 183; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 652 GTTGCCATCAACCTGATGTGACAGACATCCAGACATTTGAAATGATCATCTGCAA 711
Db 3901 GTTGCCATCAACCTGATGTGACAGACATCCAGACATTTGAAATGATCATCTGCAA 3960
QY 712 TGGCAACGAGAGGGTCCAAATGGGGCGAGCTAACAGCGACCTTTTCTGACCGAGGAGAC 771
Db 3961 TGGCAACGAGAGGGTCCAAATGGGGCGAGCTAACAGCGACCTTTTCTGACCGAGGAGAC 4020
QY 772 CACCTGGGATGTGACCTGTGGCAAAAGGTCAATTTGGAGTCCAGAGAGACCCAC 831
Db 4021 CACCTGGGATGTGACCTGTGGCAAAAGGTCAATTTGGAGTCCAGAGAGACCCAC 4080
QY 832 TGA 834
|||

Db 4081 TGA 4083
RESULT 11
ABX59427
ID ABX59427 standard; cDNA; 9732 BP.
XX
XX AC ABX59427;
XX
XX 26-FEB-2003 (first entry)
XX
XX cDNA encoding novel human musculoskeletal system antigen #1771.
DE
XX
XX Gene; ss: musculoskeletal system antigen; cancer; metastasis;
KM revascularisation; thrombosis; arteriosclerosis; mineral content;
KM cardiovascular condition; wound; injury; burn; angiogenesis; ulcer;
KM post-operative tissue repair; limb regeneration; neuronal growth;
KM neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
KM AIDS-related complex; chondrocyte growth; bone regeneration;
KM periodontal regeneration; tissue transport; bone graft; skin aging;
KM keratinocyte growth; hair loss; melanocyte growth; cell proliferation;
KM cell growth; organ transplant; cell differentiation; body height;
KM weight; hair colour; eye colour; skin; percentage of adipose tissue;
KM pigmentation; cosmetic surgery; metabolism; biorhythm; cardiac rhythm;
KM depression; tendency for violence; pain; reproductive capability;
KM hormone level; endocrine level; appetite; libido; memory; stress;
KM storage capability; fat content; lipid content; protein content;
KM carbohydrate content; vitamin content; cofactor content;
XX nutritional component.
XX
XX Homo sapiens.
XX
XX US2002147140-A1.
XX
XX 10-OCT-2002.
XX
XX 17-JAN-2001; 2001US-0764877.
XX
XX 31-JAN-2000; 2000US-179065P.
PR 04-FEB-2000; 2000US-180628P.
PR 28-JUN-2000; 2000US-214886P.
PR 07-JUL-2000; 2000US-216647P.
PR 07-JUL-2000; 2000US-216880P.
PR 11-JUL-2000; 2000US-217487P.
PR 14-JUL-2000; 2000US-217496P.
PR 14-JUL-2000; 2000US-218290P.
PR 26-JUL-2000; 2000US-220963P.
PR 26-JUL-2000; 2000US-220964P.
PR 14-AUG-2000; 2000US-224518P.
PR 14-AUG-2000; 2000US-224519P.
PR 14-AUG-2000; 2000US-225267P.
PR 14-AUG-2000; 2000US-225268P.
PR 14-AUG-2000; 2000US-225270P.
PR 14-AUG-2000; 2000US-225447P.
PR 14-AUG-2000; 2000US-225757P.
PR 14-AUG-2000; 2000US-225758P.
PR 22-AUG-2000; 2000US-226868P.
PR 30-AUG-2000; 2000US-228924P.
PR 01-SEP-2000; 2000US-229287P.
PR 01-SEP-2000; 2000US-229343P.
PR 01-SEP-2000; 2000US-229344P.
PR 01-SEP-2000; 2000US-229345P.
PR 05-SEP-2000; 2000US-229509P.
PR 05-SEP-2000; 2000US-229513P.
PR 08-SEP-2000; 2000US-231413P.
PR 21-SEP-2000; 2000US-234223P.
PR 21-SEP-2000; 2000US-234274P.
PR 25-SEP-2000; 2000US-234997P.
PR 27-SEP-2000; 2000US-235834P.
PR 29-SEP-2000; 2000US-236327P.
PR 29-SEP-2000; 2000US-236367P.
PR 29-SEP-2000; 2000US-236368P.
PR 29-SEP-2000; 2000US-236369P.

PR 29-SEP-2000; 2000US-236370P.
PR 02-OCT-2000; 2000US-236802P.
PR 02-OCT-2000; 2000US-237037P.
PR 02-OCT-2000; 2000US-237038P.
PR 02-OCT-2000; 2000US-237039P.
PR 02-OCT-2000; 2000US-237040P.
PR 13-OCT-2000; 2000US-239335P.
PR 20-OCT-2000; 2000US-240960P.
PR 20-OCT-2000; 2000US-241785P.
PR 20-OCT-2000; 2000US-241809P.
PR 01-NOV-2000; 2000US-244617P.
PR 17-NOV-2000; 2000US-249299P.
PR 08-DEC-2000; 2000US-251856P.
PR 08-DEC-2000; 2000US-251868P.
PR 08-DEC-2000; 2000US-251869P.
XX
XX (ROSE/) ROSEN C A.
PA (RUBE/) RUBEN S M.
PA (BARA/) BARASH S C.
XX
PI Rosen CA, Ruben SM, Barash SC;
XX
XX WPI; 2003-128199/12.
XX
PT Isolated nucleic acid molecules encoding musculoskeletal system
XX associated polypeptides, useful for detecting disorders, e.g. cancer -
PS Disclosure; SEQ ID NO 2804; 321pp; English.
XX
XX The invention describes an isolated nucleic acid molecule comprising a
CC sequence encoding musculoskeletal system associated polypeptides useful
CC for detecting disorders, e.g., cancer or cancer metastases, in animals
CC or humans. The nucleic acid: stimulates re-vascularisation of ischemic
CC tissues associated with conditions such as thrombosis, arteriosclerosis,
CC and other cardiovascular conditions; treats wounds due to injuries,
CC burns, post-operative tissue repair, and ulcers; stimulates angiogenesis
CC and limb regeneration; stimulates neuronal growth; can treat and prevent
CC neuronal damage occurring in certain disorders or neurodegenerative
CC conditions, such as, Alzheimer's disease, Parkinson's disease, and
CC AIDS-related complex; stimulates chondrocyte growth, thus they can be
CC used to enhance bone and periodontal regeneration and aid in tissue
CC transports or bone grafts; prevents skin aging due to sunburn by
CC stimulating keratinocyte growth; prevents hair loss, since FGF family
CC members activate hair-forming cells and promotes melanocyte growth;
CC stimulates growth and differentiation of hematopoietic cells and bone
CC marrow cells when used in combination with other cytokines; maintains
CC organs before transplantation or for supporting cell culture of primary
CC tissues; induces tissue of mesodermal origin to differentiate in early
CC embryos; increases or decreases the differentiation or proliferation of
CC embryonic stem cells, besides, hematopoietic lineage; modulates
CC mammalian characteristics, such as, body height, weight, hair colour, eye
CC colour, skin, percentage of adipose tissue, pigmentation, size, and shape
CC (e.g., cosmetic surgery); modulates mammalian metabolism; changes
CC mammal's mental state or physical state by influencing biorhythms,
CC cardiac rhythms, depression, tendency for violence, tolerance for pain,
CC reproductive capabilities, hormonal or endocrine levels, appetite,
CC libido, memory, or stress; increases or decreases storage capabilities,
CC fat content, lipid, proteins, carbohydrate, vitamins, minerals, cofactors
CC or other nutritional components. This sequence encodes a novel human
CC musculoskeletal system antigen.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from the US patent office at
CC ftp.segdata.uspto.gov/sequence.html?DocID=20020147140.
XX
SQ Sequence 9732 BP; 2063 A; 2805 C; 2561 G; 2303 T; 0 other;
Query Match 21.9%; Score 183; DB 25; Length 9732;
Best Local Similarity 100.0%; Pred. No. 5.8e-85;
Matches 183; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
652 GTTGCCATCAACGATCGAGACACATCCAGACATTCGATGGAGCATCTGCAAA 711
|||||

Db 3901 GTTGCCATCAACGATCGAGACACATCCAGACATTCGATGGAGCATCTGCAAA 3960
Qy 712 TGGCACCAGAGAGGGTTCATATGGGCGGAGCTACACGGAACCTTTTCTGAGCCAGGGGAC 771
Db 3961 TGGCACCAGAGAGGGTTCATATGGGCGGAGCTACACGGAACCTTTTCTGAGCCAGGGGAC 4020
Qy 772 CACCCCTGGGATGTGACCTCTGGCAACGGTGCACATTTGGAGTCCAGCAGACACCCAC 831
Db 4021 CACCCCTGGGATGTGACCTCTGGCAACGGTGCACATTTGGAGTCCAGCAGACACCCAC 4080
Qy 832 TGA 834
|||
Db 4081 TGA 4083
RESULT 12
AAL36440
ID AAL36440 standard; DNA; 19125 BP.
XX
XX AAL36440;
AC
XX
XX 08-JAN-2002 (first entry)
XX
XX Human musculoskeletal system related polynucleotide SEQ ID NO 2805.
XX
XX Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
XX antiallergic; hepatotropic; antidiabetic; antiinflammatory; anticancer;
XX valnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
XX cardial; gene therapy; cancer; immune disorder; cardiovascular disorder;
XX neurological disease; infection; human; secreted protein;
XX musculoskeletal system; ds.
XX
XX Homo sapiens.
XX
XX MO200155367-A1.
XX
XX 02-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US01338.
XX
XX 31-JAN-2000; 2000US-0179065.
XX 04-FEB-2000; 2000US-0180628.
XX 24-FEB-2000; 2000US-0184664.
XX 02-MAR-2000; 2000US-0186350.
XX 16-MAR-2000; 2000US-0189874.
XX 17-MAR-2000; 2000US-0190076.
XX 18-APR-2000; 2000US-0198123.
XX 19-MAY-2000; 2000US-0205515.
XX 07-JUN-2000; 2000US-0209467.
XX 28-JUN-2000; 2000US-0214886.
XX 30-JUN-2000; 2000US-0215135.
XX 07-JUL-2000; 2000US-0216647.
XX 11-JUL-2000; 2000US-0216880.
XX 11-JUL-2000; 2000US-0217487.
XX 11-JUL-2000; 2000US-0217496.
XX 14-JUL-2000; 2000US-0218290.
XX 26-JUL-2000; 2000US-0220963.
XX 26-JUL-2000; 2000US-0220964.
XX 14-AUG-2000; 2000US-0224518.
XX 14-AUG-2000; 2000US-0224519.
XX 14-AUG-2000; 2000US-0224519.
XX 14-AUG-2000; 2000US-0225213.
XX 14-AUG-2000; 2000US-0225214.
XX 14-AUG-2000; 2000US-0225266.
XX 14-AUG-2000; 2000US-0225267.
XX 14-AUG-2000; 2000US-0225268.
XX 14-AUG-2000; 2000US-0225270.
XX 14-AUG-2000; 2000US-0225477.
XX 14-AUG-2000; 2000US-0225757.
XX 14-AUG-2000; 2000US-0225758.
XX 14-AUG-2000; 2000US-0225759.
XX 18-AUG-2000; 2000US-0226279.
XX 22-AUG-2000; 2000US-0226681.
XX 22-AUG-2000; 2000US-0226686.

PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0233397.
PR 14-SEP-2000; 2000US-0233398.
PR 14-SEP-2000; 2000US-0233399.
PR 14-SEP-2000; 2000US-0233400.
PR 14-SEP-2000; 2000US-0233401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244611.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246533.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.

PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
PI WPI; 2001-451937/48.
XX
XX Isolated polypeptide for treating, preventing and/or prognosing
PT disorders related to the musculoskeletal system including
PT musculoskeletal cancers and also for testing and detection e.g.
PT diagnosis -
XX
XX
XX Example 2; SEQ ID NO 2805; 781pp + Sequence Listing; English.
XX
XX The invention relates to novel genes (AAL34669-AAL37666) and proteins
PS (AAB03087-AAB04109) associated with the musculoskeletal system useful
PS for preventing, treating or ameliorating medical conditions e.g. by
CC protein or gene therapy. The genes are isolated from a range of human
CC tissues disclosed in the specification. The nucleic acids, proteins,
CC antibodies and (ant)agonists are useful in the diagnosis, treatment
CC and prevention of: (a) cancer, e.g. breast and ovarian cancer and
CC other cancers of the adrenal gland, bone, bone marrow, breast,
CC gastrointestinal tract, liver, lung, or urogenital; (b) immune
CC disorders e.g. Addison's disease, allergies, autoimmune hemolytic
CC anemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis;
CC (c) cardiovascular disorders such as myocardial ischaemia; (d) wound
CC healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy;
CC and (f) infectious diseases such as viral, bacterial, fungal and
CC parasitic infections.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at fcp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 19125 BP; 4263 A; 5265 C; 5444 G; 4153 T; 0 other;
OY
Db 652 GTTGCATCAACTGATCTGTCAGCAGCATTCGAGGATCTGATGTCATCTGCAGAA 711
5997 GTTGCATCAACTGATCTGTCAGCAGCATTCGAGGATCTGATGTCATCTGCAGAA 6056

QY 712 TGGCACCAGAGAGGCTCAATGGGCGAGCTACCAAGCGGACCTTTTCTGAGCCAGGGGAC 771
 DB 6057 TGGCACCAGAGAGGCTCAATGGGCGAGCTACCAAGCGGACCTTTTCTGAGCCAGGGGAC 6116
 QY 772 CACCTTGGAGTGTGACCTTGGCGAA CGGTCACTTTGGAGTCCAGCAGCAGACCCAC 831
 DB 6117 CACCTTGGAGTGTGACCTTGGCGAA CGGTCACTTTGGAGTCCAGCAGCAGACCCAC 61176
 QY 832 TGA 834
 DB 6177 TGA 6179

RESULT 13
 ABX59428
 ID ABX59428 standard; cDNA; 19125 BP.
 AC ABX59428;
 XX
 DT 26-FEB-2003 (first entry)
 DE cDNA encoding novel human musculoskeletal system antigen #1772.
 XX
 KM Gene; ss; musculoskeletal system antigen; cancer; metastasis;
 KM re-vascularisation; thrombosis; arteriosclerosis; mineral content;
 KM post-operative tissue repair; limb regeneration; neuronal growth;
 KM neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
 KM AIDS-related complex; chondrocyte growth; bone regeneration;
 KM periodontal regeneration; tissue transport; bone graft; skin aging;
 KM keratinocyte growth; hair loss; melanocyte growth; cell proliferation;
 KM cell growth; organ transplant; cell differentiation; body height;
 KM weight; hair colour; eye colour; skin; percentage of adipose tissue;
 KM pigmentation; cosmetic surgery; metabolism; biorthym; cardiac rhythm;
 KM depression; tendency for violence; pain; reproductive capability;
 KM hormone level; endocrine level; appetite; libido; memory; stress;
 KM storage capability; fat content; lipid content; protein content;
 KM carbohydrate content; vitamin content; cofactor content;
 KM nutritional component.
 XX
 OS Homo sapiens.
 XX
 PN US2002147140-A1.
 XX
 PD 10-OCT-2002.
 XX
 PF 17-JAN-2001; 2001US-0764877.
 XX
 PR 31-JAN-2000; 2000US-179065P.
 PR 04-FEB-2000; 2000US-180628P.
 PR 28-JUN-2000; 2000US-214886P.
 PR 07-JUL-2000; 2000US-216647P.
 PR 07-JUL-2000; 2000US-216880P.
 PR 11-JUL-2000; 2000US-217487P.
 PR 11-JUL-2000; 2000US-217496P.
 PR 14-JUL-2000; 2000US-218290P.
 PR 26-JUL-2000; 2000US-220963P.
 PR 26-JUL-2000; 2000US-220964P.
 PR 14-AUG-2000; 2000US-224518P.
 PR 14-AUG-2000; 2000US-224519P.
 PR 14-AUG-2000; 2000US-225267P.
 PR 14-AUG-2000; 2000US-225268P.
 PR 14-AUG-2000; 2000US-225270P.
 PR 14-AUG-2000; 2000US-225447P.
 PR 14-AUG-2000; 2000US-225757P.
 PR 14-AUG-2000; 2000US-225758P.
 PR 22-AUG-2000; 2000US-226868P.
 PR 30-AUG-2000; 2000US-228924P.
 PR 01-SEP-2000; 2000US-229287P.
 PR 01-SEP-2000; 2000US-229343P.
 PR 01-SEP-2000; 2000US-229344P.
 PR 01-SEP-2000; 2000US-229345P.

PR 05-SEP-2000; 2000US-229509P.
 PR 05-SEP-2000; 2000US-229513P.
 PR 08-SEP-2000; 2000US-231413P.
 PR 21-SEP-2000; 2000US-234223P.
 PR 21-SEP-2000; 2000US-234274P.
 PR 25-SEP-2000; 2000US-234997P.
 PR 27-SEP-2000; 2000US-235834P.
 PR 29-SEP-2000; 2000US-236337P.
 PR 29-SEP-2000; 2000US-236357P.
 PR 29-SEP-2000; 2000US-236368P.
 PR 29-SEP-2000; 2000US-236369P.
 PR 29-SEP-2000; 2000US-236370P.
 PR 02-OCT-2000; 2000US-236802P.
 PR 02-OCT-2000; 2000US-237037P.
 PR 02-OCT-2000; 2000US-237038P.
 PR 02-OCT-2000; 2000US-237039P.
 PR 02-OCT-2000; 2000US-237040P.
 PR 13-OCT-2000; 2000US-237935P.
 PR 20-OCT-2000; 2000US-240960P.
 PR 20-OCT-2000; 2000US-241785P.
 PR 20-OCT-2000; 2000US-241809P.
 PR 01-NOV-2000; 2000US-244617P.
 PR 17-NOV-2000; 2000US-249299P.
 PR 08-DEC-2000; 2000US-251856P.
 PR 08-DEC-2000; 2000US-251868P.
 PR 08-DEC-2000; 2000US-251869P.
 XX
 PA (ROSE/) ROSEN C A.
 PA (RUBE/) RUBE S M.
 PA (BARA/) BARASH S C.
 XX
 PI Rosen CA, Ruben SM, Barash SC;
 DR WPI; 2003-128199/12.
 XX
 PT Isolated nucleic acid molecules encoding musculoskeletal system
 associated polypeptides, useful for detecting disorders, e.g. cancer -
 XX
 PS Disclosure, SEQ ID NO 2805; 321pp; English.
 XX
 CC The invention describes an isolated nucleic acid molecule comprising a
 CC sequence encoding musculoskeletal system associated polypeptides useful
 CC for detecting disorders, e.g., cancer or cancer metastases, in animals
 CC or humans. The nucleic acid stimulates re-vascularisation of ischaemic
 CC tissues associated with conditions such as thrombosis, arteriosclerosis,
 CC and other cardiovascular conditions; treats wounds due to injuries,
 CC burns, post-operative tissue repair, and ulcers; stimulates angiogenesis
 CC and limb regeneration; stimulates neuronal growth; can treat and prevent
 CC neuronal damage occurring in certain disorders or neurodegenerative
 CC conditions, such as, Alzheimer's disease, Parkinson's disease, and
 CC AIDS-related complex; stimulates chondrocyte growth, thus they can be
 CC used to enhance bone and periodontal regeneration and aid in tissue
 CC transports or bone grafts; prevents skin aging due to sunburn by
 CC stimulating keratinocyte growth; prevents hair loss, since FGF family
 CC members activate hair-forming cells and promotes melanocyte growth;
 CC stimulates growth and differentiation of hematopoietic cells and bone
 CC marrow cells when used in combination with other cytokines; maintains
 CC organs before transplantation or for supporting cell culture of primary
 CC tissues; induces tissue of mesodermal origin to differentiate in early
 CC embryonic; increases or decreases the differentiation or proliferation of
 CC embryonic stem cells, besides, hematopoietic lineage; modulates
 CC mammalian characteristics, such as, body height, weight, hair colour, eye
 CC colour, skin, percentage of adipose tissue, pigmentation, size, and shape
 CC (e.g., cosmetic surgery); modulates mammalian metabolism; changes
 CC mammal's metal state or physical state by influencing biorthym; changes
 CC cardiac rhythms, depression, tendency for violence, tolerance for pain,
 CC reproductive capabilities, hormonal or endocrine levels, appetite,
 CC libido, memory or stress; increases or decreases storage capabilities,
 CC fat content, lipid, protein, carbohydrate, vitamin, minerals, cofactors
 CC or other nutritional components. This sequence encodes a novel human
 CC musculoskeletal system antigen.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly

CC from the US patent office at
CC ftp://seqdata.uspto.gov/sequence.html?docid=20020147140.
XX
SQ Sequence 19125 BP; 4263 A; 5265 C; 5444 G; 4153 T; 0 other;
Query Match 21.9%; Score 183; DB 25; Length 19125;
Best Local Similarity 100.0%; Pred. No. 5.9e-85;
Matches 183; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 652 GTTGCATCAACCTGATGTCGACACATCCAGACATTCGATGTCGACATCTGCAAA 711
DB 5997 GTTGCATCAACCTGATGTCGACACATCCAGACATTCGATGTCGACATCTGCAAA 6056
QY 712 TGGCACCAGGAGGGTCCAAATGCGCGAGCTACTAAAGCGAGCTTTCTGACCCAGGGGAC 771
DB 6057 TGGCACCAGGAGGGTCCAAATGCGCGAGCTACTAAAGCGAGCTTTCTGACCCAGGGGAC 6116
QY 772 CACCTTGGGATGCTGACCTTGGCAAAAGGTCACATTTGAGTCCAGCAGACACCCGAC 831
DB 6117 CACCTTGGGATGCTGACCTTGGCAAAAGGTCACATTTGAGTCCAGCAGACACCCGAC 6176
QY 832 TGA 834
DB 6177 TGA 6179
RESULT 14
ABN51411
ID ABN51411 standard; DNA; 65 BP.
AC ABN51411;
XX
DT 15-JUL-2002 (first entry)
XX
DE Mouse spliced transcript detection oligonucleotide SEQ ID NO:24159.
XX
KW Human; mouse; rat; splice transcript; detection; RNA transcript;
KW splice variant; transcriptome; oligonucleotide library; ss.
XX
OS Mus musculus.
XX
FN WO200210449-A2.
XX
PD 07-FEB-2002.
XX
PF 20-JUL-2001; 2001WO-1B01903.
XX
PR 28-JUL-2000; 2000US-221607P.
PR 02-MAY-2001; 2001US-287724P.
PA (COMP-) COMPUGEN INC.
XX
PI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S,
XX
DR WPI; 2002-257383/30.
XX
PT New oligonucleotide libraries comprising oligonucleotides which
PT selectively hybridize to mRNAs transcribed from a transcription unit of
PT a genome, useful for detecting tissue-, pathology-, and
PT developmental-specific genes.
XX
PS Example 1; SEQ ID 24159; 47bp; English.
XX
CC The present invention describes oligonucleotide libraries for detecting
CC messenger RNAs that populate a (sub-)transcriptome, where the
CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC transcription units that populate a genome. The library comprises
CC several oligonucleotides, each capable of hybridizing selectively to a
CC set of messenger RNAs transcribed from a given transcription unit of
CC the genome, which encodes one or more messenger RNA splice variants.
CC The oligonucleotide libraries are useful for detecting mRNAs from a
CC biological sample, in expression profiling studies, in qualitatively or
CC quantitatively characterizing the corresponding transcriptome, and in

CC detecting RNA transcripts and splice variants of human or animal
CC transcriptomes. The libraries may also be used as specialised mini
CC libraries to detect transcripts of a sub-transcriptome under a
CC particular biological or pathological state, and so allowing the
CC detection of tissue- and pathology-specific genes such as those genes
CC only expressed in specific tissue under a specific pathological
CC condition; to detect developmental specific genes; and to detect RNA
CC transcripts and splice variants of a transcriptome of a patient suffering
CC from a particular disorder. ABN27253 to ABN59589 represent
CC oligonucleotide sequences from rats, humans and mice, which are used in
CC the exemplification of the present invention.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp://wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 65 BP; 16 A; 13 C; 22 G; 14 T; 0 other;
Query Match 3.1%; Score 26; DB 24; Length 65;
Best Local Similarity 100.0%; Pred. No. 0.0043;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 493 AGGCTGTCTGGAAGAGTCTCCGGGA 518
DB 16 AGGCTGTCTGGAAGAGTCTCCGGGA 41
RESULT 15
AAZ15454
ID AAZ15454 standard; cDNA; 715 BP.
AC AAZ15454;
XX
DT 12-OCT-1999 (first entry)
XX
DE Human gene expression product cDNA sequence SEQ ID NO:2923.
XX
KW Human; gene; gene expression product; diagnosis; therapy; probe;
KW detection; mapping; tissue typing; profiling; forensic; cancer;
KW genetic analysis; colorectal cancer; breast cancer; lung cancer; ss.
XX
OS Homo sapiens.
XX
FN WO9938972-A2.
XX
PD 05-AUG-1999.
XX
PF 28-JAN-1999; 99WO-US01619.
XX
PR 03-APR-1998; 98US-0080666.
PR 28-JAN-1998; 98US-0072910.
PR 24-FEB-1998; 98US-0075954.
PR 31-MAR-1998; 98US-0080114.
PR 03-APR-1998; 98US-0080515.
XX
PA (CHIR) CHIRON CORP.
PA (HYSE-) HYSEQ INC.
XX
PI Ctkvenjakov R, Dickson M, Dymnac R, Dymnac S;
PI Escobedo J, Garcia PD, Garcia V, Glese K, Imils MA;
PI Jones WL, Kassam A, Kennedy GC, Kita D, Labat I;
PI Lamson G, Leshkowitz D, Pot D, Randazzo F, Reinhard C;
PI Stache-Crain B, Suduth-Klinger J, Williams LR;
XX
DR WPI; 1999-494092/41.
XX
PT Novel human genes and their expression products which are
PT differentially expressed in different cell types
XX
PS Claim 1; Page 1410-1411; 2479pp; English.
XX
CC The present invention describes a library of human polynucleotides
CC comprising the sequences given in AAZ12532 to AAZ1779. Also described is
CC a method of detecting differentially expressed genes correlated with the

CC cancerous state of a mammalian cell, comprising detecting at least one
CC differentially expressed gene product in a test sample from a cell
CC suspected of being cancerous, where the gene product is encoded by one
CC of the 5248 polynucleotide sequences given in AA212532 to AA217779. The
CC polynucleotides can be used as a source of primers and probes, which can
CC be used for a variety of purpose, e.g. detection of expression levels,
CC mapping, tissue typing or profiling, forensics, genetic analysis and
CC detection of polymorphisms. Polypeptides encoded by the polynucleotides
CC can be used for raising antibodies for experimental, diagnostic and
CC therapeutic purposes. The polynucleotides may also be used to construct
CC arrays for diagnostics (which may be used to determine function of an
CC encoded protein); and to detect differences in expression levels between
CC two cells (e.g. to identify abnormal or diseased tissue in a human, to
CC identify a genetic predisposition or susceptibility to a disease such as
CC cancer). The polynucleotides of the invention are especially used in the
CC diagnosis, prognosis and management of colorectal cancer, breast cancer,
CC and lung cancer. The polynucleotides can also be used to screen for
CC peptide analogues and antagonists.

XX
SQ Sequence 715 BP, 179 A, 189 C, 178 G, 157 T, 12 other;

Query Match 2.8%; Score 23; DB 20; Length 715;

Best Local Similarity 100.0%; Pred. No. 0.17;

Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 667 ATCGTGCAGCATCCAGACAT 689
DB 572 ATCGTGCAGCATCCAGACAT 594

Search completed: November 25, 2003, 02:31:21
Job time : 258.154 secs

This Page Blank (uspto)

PR 19-OCT-2000; 2000US-0693036.
 PR 29-NOV-2000; 2000US-0727344.
 XX
 XX (HYSE-) HYSEQ INC.
 PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
 PI Zhao QA, Zhou F, Goodrich R, Drmanac RT;
 XX
 DR WPI; 2001-442253/47.
 DR P-PSDB; AAM39502.
 PT Novel nucleic acids and polypeptides, useful for treating disorders
 PT such as central nervous system injuries -
 PS
 PS Claim 1; SEQ ID NO 861; 10078bp; English.
 XX
 XX The invention relates to human nucleic acids (AA157798-AA161369) and
 CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localized neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilization of the activities such as: Immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukemias and
 CC C.N.S disorders
 CC Note: the sequence data for this patent did not form part of the printed
 CC specification.
 CC
 SQ Sequence 1288 BP; 275 A; 361 C; 418 G; 234 T; 0 other;
 XX
 Query Match 100.0%; Score 834; DB 22; Length 1288;
 Best Local Similarity 100.0%; Pred. No. 8,2e-206;
 Matches 834; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ATGGCTTCGGCGGAGGCGGAAGACTGCGAGAGCCCGCGCGCGAGCCGTCGCCGAC 60
 DB 95 ATGGCTTCGGCGGAGGCGGAAGACTGCGAGAGCCCGCGCGCGAGCCGTCGCCGAC 154
 QY 61 CAGCGGCCCTCTCTGATGAGGGGTGACCGCGGCACTGCGACCGGGAAGTCACTGTGT 120
 DB 155 CAGCGGCCCTCTCTGATGAGGGGTGACCGCGGCACTGCGACCGGGAAGTCACTGTGT 214
 QY 121 GAGAAGATCATGAGTGTCTGGGACAGAAAGAGGTGAAACGCGGACCGGGAAGTGTGTC 180
 DB 215 GAGAAGATCATGAGTGTCTGGGACAGAAAGAGGTGAAACGCGGACCGGGAAGTGTGTC 274
 QY 181 ATCTGAGCCAGGACAGGTTCTACAAAGTCTGACGCGAGAGCAAAAGCCAGGCTTGG 240
 DB 275 ATCTGAGCCAGGACAGGTTCTACAAAGTCTGACGCGAGAGCAAAAGCCAGGCTTGG 334
 QY 241 AAAGACAGTACATTTTGAACATCCAGATGCTTTGATATATGATTTGATGACAGACT 300
 DB 335 AAAGACAGTACATTTTGAACATCCAGATGCTTTGATATATGATTTGATGACAGACT 394
 QY 301 CTGAAGAACATCGTGAAGGCGAAACGCGTGAAGGCGGACCTATGATTTTGTGACACAC 360
 DB 395 CTGAAGAACATCGTGAAGGCGAAACGCGTGAAGGCGGACCTATGATTTTGTGACACAC 454
 QY 361 TCAAGGTTACAGAGACCAAGCGGTGTACCTCGGAGCGTGTGTTGTTGAGGGCATC 420
 DB 455 TCAAGGTTACAGAGACCAAGCGGTGTACCTCGGAGCGTGTGTTGTTGAGGGCATC 514
 QY 421 TTGGTGTCTACAGCGAGAGATCCGGGACATGTTTCCACTGCGGCTTTTGTGTGACACC 480
 DB 515 TTGGTGTCTACAGCGAGAGATCCGGGACATGTTTCCACTGCGGCTTTTGTGTGACACC 574
 QY 481 GACTCCGACGTCAGGCTGTCTCGAAGAGTTCTCCGGGAGTGGCGGAGGAGGAGCACTTG 540

DB 575 GACTCCGACGTCAGGCTGTCTCGAAGAGTTCTCCGGAGAGTGGCGGAGGAGGACCTTG 634
 QY 541 GAGCAGATTCTGACGCGATACACACACTTTCGTGAACCGGCTTGGAGAGTTCTGCTTG 600
 DB 635 GAGCAGATTCTGACGCGATACACACACTTTCGTGAACCGGCTTGGAGAGTTCTGCTTG 694
 QY 601 CCGACAAAGAGTATGCCATGTGATCATCCACGAGAGTGGACAAATGTTGTCATC 660
 DB 695 CCGACAAAGAGTATGCCATGTGATCATCCACGAGAGTGGACAAATGTTGTCATC 754
 QY 661 AACCTGATGTCAGACATCCAGACGACATTTGAAATGTGATCATGCAATGACACCGA 720
 DB 755 AACCTGATGTCAGACATCCAGACGACATTTGAAATGTGATCATGCAATGACACCGA 814
 QY 721 GGAGGTCGAATGGCGGAGCTACAAACGCGACCTTTCTGAGCCAGGAGACCACTTGG 780
 DB 815 GGAGGTCGAATGGCGGAGCTACAAACGCGACCTTTCTGAGCCAGGAGACCACTTGG 874
 QY 781 ATGCTGACCTTCGGGCAACGCTCACATTTGAGTCCAGACGAGACCCGACTGA 834
 DB 875 ATGCTGACCTTCGGGCAACGCTCACATTTGAGTCCAGACGAGACCCGACTGA 928
 RESULT 2
 AAD27186
 ID AAD27186 standard; DNA; 1624 BP.
 XX
 XX AAD27186;
 AC
 XX
 DT 18-APR-2002 (first entry)
 XX
 DE Human 57658 DNA.
 XX
 XX Human; uridine kinase-like protein; haematopoietic neoplastic disorder;
 KW 57658 protein; leukaemia; haematopoietic disorder; diabetes mellitus;
 KW rheumatoid arthritis; psoriasis; dermatitis; osteoporosis; rickets;
 KW sarcoma; myocardial infarction; hypertension; atherosclerosis;
 KW Alzheimer's disease; Parkinson's disease; Huntington's disease;
 KW gene therapy; chromosome mapping; tissue typing; dermatological;
 KW cytostatic; osteopathic; cardiant; neuroprotective; nootropic;
 KW anticonvulsant; ds.
 KW
 OS Homo sapiens.
 XX
 XX
 FH Key Location/Qualifiers
 FT CDS 94..927
 FT /*tag= a
 FT /product= "Human 57658 protein"
 FT misc_feature 94..924
 FT /*tag= b
 FT /note= "This region is specifically claimed as
 FT SEQ ID NO: 3 in claim 1 of the specification"
 FT
 PD WO200202761-A2.
 XX
 PD 10-JAN-2002.
 XX
 XX 28-JUN-2001; 2001WO-US21063.
 PF
 XX 30-JUN-2000; 2000US-216503P.
 PR
 PA (MILL-) MILLENNIUM PHARM INC.
 XX
 XX Glucksmann MA;
 PI WPI; 2002-140091/18.
 DR P-PSDB; AAB16592.
 XX
 PT New isolated human uridine kinase family polypeptide 57658, useful for
 PT treating hematopoietic neoplastic disorders and disorders of neurons,
 PT heart and blood vessels -
 PT
 XX

PS Claim 1; Fig 1a; 103bp; English.

XX The patent discloses human uridine kinase-like polypeptides, designated
CC 57658 and polynucleotides encoding such proteins. 57658 DNAs and proteins
CC are useful for developing diagnostic and therapeutic agents for 57658-
CC mediated or related disorders such as haematopoietic neoplastic disorders
CC (e.g. leukaemia), haematopoietic disorders (e.g. psoriasis, dermatitis,
CC diabetes mellitus, rheumatoid arthritis, disorders of bone (e.g. rickets,
CC osteoporosis), heart (e.g. sarcoma, myocardial infarction), blood vessels
CC (e.g. hypertension, atherosclerosis) and neurons (e.g. Huntington's
CC disease, Alzheimer's disease, Parkinson's disease). 57658 sequences or
CC their antibodies are useful in screening assays, detection assays (e.g.
CC forensic biology) and predictive medicine (e.g. diagnostic assays).
CC prognostic assays, and monitoring clinical trials and pharmacogenomics).
CC They are useful as reagents for diagnosing and treating 57658-mediated
CC disorders. 57658 DNAs are useful in gene therapy, in chromosome mapping,
CC to identify an individual from a minute biological sample (tissue typing)
CC and to aid in forensic identification of the biological sample. The
CC present sequence is a DNA encoding human 57658 protein.

XX Sequence 1624 BP; 354 A; 427 C; 506 G; 337 T; 0 other;

Query Match 100.0%; Score 834; DB 24; Length 1624;

Best Local Similarity 100.0%; Pred. No. 8.9e-206;

Matches 834; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCTTCGGCGGAGGCGAAGACTGCGAAGACCCCGCGGAGGCGGACCGTCCGAC 60
DB ATGGCTTCGGCGGAGGCGAAGACTGCGAAGACCCCGCGGAGGCGGACCGTCCGAC 153
QY 61 CAGCGGCGCTTCTGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 120
DB 154 CAGCGGCGCTTCTGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 213
QY 121 GAGAAGATCATGAGTGTCTGCGGACAGAACGAGGTGAGACGCGGACGCGAAGTGTGTC 180
DB 214 GAGAAGATCATGAGTGTCTGCGGACAGAACGAGGTGAGACGCGGACGCGAAGTGTGTC 273
QY 181 ATCTGAGCCAGGACGAGTCTTACAGAGTCTTACAGGACGAGGAGGAGGAGGAGGAGG 240
DB 274 ATCTGAGCCAGGACGAGTCTTACAGAGTCTTACAGGACGAGGAGGAGGAGGAGGAGG 333
QY 241 AAGGACGATGACATTTTGAACATCCAGATGCGCTTATATATATTTGATGACAGGACT 300
DB 334 AAGGACGATGACATTTTGAACATCCAGATGCGCTTATATATATTTGATGACAGGACT 393
QY 301 CTGAGAACATCGTGAAGGCGAAGACGAGTGGCGGACCTATGATTTTGTGACACAC 360
DB 394 CTGAGAACATCGTGAAGGCGAAGACGAGTGGCGGACCTATGATTTTGTGACACAC 453
QY 361 TCAAGTTTACAGAGACACGAGTGTCTTACCTGCGGACGAGTGTCTGTTGAGGAGATC 420
DB 454 TCAAGTTTACAGAGACACGAGTGTCTTACCTGCGGACGAGTGTCTGTTGAGGAGATC 513
QY 421 TTGAGTTTACAGGACGAGGATCCGCGGACATGTTCCACTGGCGCTCTCGAGACAC 480
DB 514 TTGAGTTTACAGGACGAGGATCCGCGGACATGTTCCACTGGCGCTCTCGAGACAC 573
QY 481 GACTCCGACGTGAGGCTGTCTCGAAGAGTCTCCGAGACGTGGCGGACGAGGAGGAGCTG 540
DB 574 GACTCCGACGTGAGGCTGTCTCGAAGAGTCTCCGAGACGTGGCGGACGAGGAGGAGCTG 633
QY 541 GACGAGATTTGAGCGGACGACCACTTGTGTAAGCGGCGCTTCCAGGAGTGTCTGCTG 600
DB 634 GACGAGATTTGAGCGGACGACCACTTGTGTAAGCGGCGCTTCCAGGAGTGTCTGCTG 693
QY 601 CCGACAAAGAGATGCGGATGATCATCCACGAGAGTGGACATATGTTGGCATC 660
DB 694 CCGACAAAGAGATGCGGATGATCATCCACGAGAGTGGACATATGTTGGCATC 753
QY 661 AACCTGATGTGAGCAGCATCCAGAGATTTCTGATGTGACATCTTGCAGAAATGGACCGA 720
DB 754 AACCTGATGTGAGCAGCATCCAGAGATTTCTGATGTGACATCTTGCAGAAATGGACCGA 813

QY 721 GGAGGGTCCAATGAGGAGGAGCTTACAGCGGACCTTTTCTGAGCCAGGAGGACACCTGGG 780
DB 814 GGAGGGTCCAATGAGGAGGAGCTTACAGCGGACCTTTTCTGAGCCAGGAGGACACCTGGG 873
QY 781 ATCTGACCTCTTGCGAAGCGGTGCATTTTGAATCCAGAGCAGACCCCACTGA 834
DB 874 ATCTGACCTCTTGCGAAGCGGTGCATTTTGAATCCAGAGCAGACCCCACTGA 927

RESULT 3

AAH15621
ID AAH15621 strand; cDNA; 2160 BP.

AC AAH15621;

DT 26-JUN-2001 (first entry)

DE Human cDNA sequence SEQ ID NO:13951.

KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

OS Homo sapiens.

PN EP1074617-A2.

PD 07-FEB-2001.

PF 28-JUL-2000; 2000EP-0116126.

PR 29-JUL-1999; 99JP-0248036.

PR 27-AUG-1999; 99JP-0300253.

PR 11-JUN-2000; 2000JP-0118776.

PR 02-MAY-2000; 2000JP-0183767.

PR 09-JUN-2000; 2000JP-0241899.

PA (HELT-) HELIX RES INSTR.

PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

DR WPI; 2001-318749/34.

PS Claim 8; SEQ ID 13951; 2537pp + CD ROM; English.

XX The present invention describes primer sets for synthesizing 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH1633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH1632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.

SQ Sequence 2160 BP; 457 A; 591 C; 671 G; 441 T; 0 other;
 Query Match 100.0%; Score 834; DB 22; Length 2160;
 Best Local Similarity 100.0%; Pred. No. 9.8e-206;
 Matches 834; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 ATGGCTTCGGGGGGGAGGCGAAGACTGCGAGAGCCCGGCGGAGGCGGACCGCGCAC 60
DB 95 ATGGCTTCGGGGGGGAGGCGAAGACTGCGAGAGCCCGGCGGAGGCGGACCGCGCAC 154
QY 61 CAGCGGCCCTTCTGATAGGGGTGAGCGCGGCACTGCGAGCGGAAAGTCAGCCGTGTGT 120
DB 155 CAGCGGCCCTTCTGATAGGGGTGAGCGCGGCACTGCGAGCGGAAAGTCAGCCGTGTGT 214
QY 121 GAGAGATCATGAGTTGCTGGGACAGAACGAGGTGAAACGCGGACAGCGGAAGTGTGTC 180
DB 215 GAGAGATCATGAGTTGCTGGGACAGAACGAGGTGAAACGCGGACAGCGGAAGTGTGTC 274
QY 181 ATCTGAGCCGAGCAGGTTCTACAGGTTCTGACGGGAGAGCAGAAAGGCCAGGCTTG 240
DB 275 ATCTGAGCCGAGCAGGTTCTACAGGTTCTGACGGGAGAGCAGAAAGGCCAGGCTTG 334
QY 241 AAAGCAGATCAATTTTGAACATCCAGATGCTTTGATATGATTTGATCAGAGACT 300
DB 335 AAAGCAGATCAATTTTGAACATCCAGATGCTTTGATATGATTTGATCAGAGACT 394
QY 301 CTGAAGAACATGCGAGAGGCGAAACGCTGAGAGTCCGACCTATGATTTGTGACACAC 360
DB 395 CTGAAGAACATGCGAGAGGCGAAACGCTGAGAGTCCGACCTATGATTTGTGACACAC 454
QY 361 TCAGAGTTACAGAGACCCAGGTTGTCTACCCCTGCGAGCGTGTCTGTTGAGGGATC 420
DB 455 TCAGAGTTACAGAGACCCAGGTTGTCTACCCCTGCGAGCGTGTCTGTTGAGGGATC 514
QY 421 TTGGTGTCTACAGCCAGAGATCCGGGACATGTTCCACCTGCGCTCTTGTGTGACACC 480
DB 515 TTGGTGTCTACAGCCAGAGATCCGGGACATGTTCCACCTGCGCTCTTGTGTGACACC 574
QY 481 GACTCCGACGTCAGGCTGTCTCGAAGAGTTCTCCGGGACGTGCGCCGAGGAGGAGACTG 540
DB 575 GACTCCGACGTCAGGCTGTCTCGAAGAGTTCTCCGGGACGTGCGCCGAGGAGGAGACTG 634
QY 541 GAGAGATTTGAGCGAGATCACCACTTCGTAAGCCGCGCTTCGAGAGATTTGTGCTG 600
DB 635 GAGAGATTTGAGCGAGATCACCACTTCGTAAGCCGCGCTTCGAGAGATTTGTGCTG 694
QY 601 CCGACAAAGAGATGCGGATGTATCATCCACGAGAGTGAACAATATGTTGCCATC 660
DB 695 CCGACAAAGAGATGCGGATGTATCATCCACGAGAGTGAACAATATGTTGCCATC 754
QY 661 AACCTGATCGTGCAGCATTCAGAGCACTTCTGAATGTGTGCAATCTGCAATGGCACCGA 720
DB 755 AACCTGATCGTGCAGCATTCAGAGCACTTCTGAATGTGTGCAATCTGCAATGGCACCGA 814
QY 721 GGAGGGTTCATGAGCGGAGCTACAAAGCGGACCTTTTGTGCGCAGGGGACACCCCTGGG 780
DB 815 GGAGGGTTCATGAGCGGAGCTACAAAGCGGACCTTTTGTGCGCAGGGGACACCCCTGGG 874
QY 781 ATGTGACCTCTGGCAACGCTCATTTGGAGTTCAGACAGACCCCACTGA 834
DB 875 ATGTGACCTCTGGCAACGCTCATTTGGAGTTCAGACAGACCCCACTGA 928
  
```

RESULT 4
 ID AAK98735
 AA98735 standard; DNA; 834 BP.

AC AAK98735;
 DT 02-MAY-2002 (first entry)
 DE DNA of a human uridine kinase (UDK).

KW Human; uridine kinase; diagnostic assay; mutation detection; UDK;
 KW probe; chromosome localisation study; tissue expression; gene therapy;
 KW antibody; vaccine; human ovarian cancer; immunological disorder;
 KW human colon carcinoma; immunogen; ds.
 OS Homo sapiens.

XX	Key	Location/Qualifiers
FT	CDS	1..780
FT		/tag= a
FT		/partial
XX		/note= "No stop codon"
PN		MO200172963-A2.
XX		04-OCT-2001.
PD		
XX		27-MAR-2001; 2001MO-US09663.
PF		27-MAR-2000; 2000US-0536647.
XX		
PA		(SMIK) SMITHKLINE BEECHAM CORP.
PI		Ho YS, Johnson RK;
XX		
DR		WPI, 2001-626259/72.
XX		P-PSDB; AAO14412.
PT		Novel human uridine kinase polypeptides useful for treating cancers,
PT		and to identify agonists and antagonists of the polypeptide useful for
PT		treating conditions associated with uridine kinase imbalance
PS		Claim 5; Page 22-23; 31pp; English.
XX		

The invention relates to newly identified human uridine kinase (UDK)
 polypeptides and polynucleotides and methods for producing such
 polypeptides by recombinant techniques. Also disclosed in the invention
 are methods for utilising uridine kinase polypeptides and polynucleotides
 in diagnostic assays. The polynucleotides and polypeptides of the
 invention may be used as diagnostic reagents by detecting mutations in an
 associated gene. An array of oligonucleotide probes comprising the
 uridine kinase polynucleotide sequence or fragments thereof can be
 constructed to conduct efficient screening of genetic mutations, for
 example, detection of abnormally decreased or increased levels of
 polypeptide or mRNA expression may also be used for diagnosing or
 determining susceptibility of a subject to a disease of the invention.
 The polynucleotide sequences of the invention can be used for chromosome
 localisation studies and tissue expression studies. The polypeptides of
 the invention or fragments thereof may be used as immunogens to produce
 antibodies. These antibodies may be employed to isolate or identify
 clones expressing the polypeptide. The polypeptides and polynucleotides
 of the invention can be used as a vaccine or in gene therapy to treat
 diseases such as human ovarian cancer, human colon carcinomas, and
 immunological disorders. This polynucleotide sequence represents the DNA
 of a human uridine kinase of the invention.

SQ Sequence 834 BP; 194 A; 219 C; 265 G; 156 T; 0 other;

Query Match 99.8%; Score 832.4; DB 23; Length 834;
 Best Local Similarity 99.9%; Pred. No. 1.8e-205;
 Matches 833; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 ATGGCTTCGGGGGGGAGGCGAAGACTGCGAGAGCCCGGCGGAGGCGGACCGCTCCGCAC 60
DB 1 ATGGCTTCGGGGGGGAGGCGAAGACTGCGAGAGCCCGGCGGAGGCGGACCGCTCCGCAC 60
QY 61 CAGCGGCCCTTCTGATAGGGGTGAGCGGCGGCACTGCGAGCGGAAAGTGCACCGTGTGT 120
DB 61 CAGCGGCCCTTCTGATAGGGGTGAGCGGCGGCACTGCGAGCGGAAAGTGCACCGTGTGT 120
QY 121 GAGAGATCATGAGTTGCTGGGACAGAACGAGGTGAAACGCGGACAGCGGAAGTGTGTC 180
DB 121 GAGAGATCATGAGTTGCTGGGACAGAACGAGGTGAAACGCGGACAGCGGAAGTGTGTC 180
  
```


QY	181	ATCCGACCCAGSAGAGGTTCTACAAAGTCCGACGCGAGAGAGAAAGCCAAAGCCCTTG	240
Db	181	ATCCGAGCAGAGACAGGTTCTACAAAGTCTGACCGCAGAGAGAAAGCCAAAGCCCTTG	240
QY	241	AAAGGACAGTAAATTTTGAACCATCCAGATGCGCTTGTATTAATGATTTGATGACAGAACT	300
Db	241	AAAGGACAGTAAATTTTGAACCATCCAGATGCGCTTGTATTAATGATTTGATGACAGAACT	300
QY	301	CTGAAGAACATCGTGAGGGCAAAAACGCTGAGAGGTCCGACCTATGATTTTGTGACACAC	360
Db	301	CTGAAGAACATCGTGAGGGCAAAAACGCTGAGAGGTCCGACCTATGATTTTGTGACACAC	360
QY	361	TCAAGGTACCAAGACCAACGGTGGCTACCCCTGCGSAGAGTGTCTGTTTGAAGGGCAATC	420
Db	361	TCAAGGTACCAAGACCAACGGTGGCTACCCCTGCGSAGAGTGTCTGTTTGAAGGGCAATC	420
QY	421	TTGCTGTTCTACAGCCAGAGAGATCCGGGACATGTTCCACTGGCGCTCTTCCGACACAC	480
Db	421	TTGCTGTTCTACAGCCAGAGAGATCCGGGACATGTTCCACTGGCGCTCTTCCGACACAC	480
QY	481	GACTCCGACGTCAAGGCTGTCTCGAAGAGTTCCTCGGGAAGTGGCCGAGGAGGACCTG	540
Db	481	GACTCCGACGTCAAGGCTGTCTCGAAGAGTTCCTCGGGAAGTGGCCGAGGAGGACCTG	540
QY	541	GAGCAGATTTCTGACCGCAGTACACACCTTCGTGAGACCGGCTTCCAGAGAGTTCTGCTG	600
Db	541	GAGCAGATTTCTGACCGCAGTACACACCTTCGTGAGACCGGCTTCCAGAGAGTTCTGCTG	600
QY	601	CCGACAAAGAAAGTATGCCGATGTGATCATCCACGAGAGTGGACAAATAGTTGGCAATC	660
Db	601	CCGACAAAGAAAGTATGCCGATGTGATCATCCACGAGAGTGGACAAATAGTTGGCAATC	660
QY	661	AACTGATGCTGACAGACATCCAGGACATTTCTGAATGTTGACATCTTGCAAAATGGACCGA	720
Db	661	AACTGATGCTGACAGACATCCAGGACATTTCTGAATGTTGACATCTTGCAAAATGGACCGA	720
QY	721	GGAGGGTCCAAATGGGGCGAGCTACAAAGCGGACCTTTTCTGAGCCAGGGGACCAACCTGGG	780
Db	721	GGAGGGTCCAAATGGGGCGAGCTACAAAGCGGACCTTTTCTGAGCCAGGGGACCAACCTGGG	780
QY	781	ATGCTGACCTCTGGCAAAACGCTACATTTTGGAGTCCAGACGACGACCCCACTGA	834
Db	781	ATGCTGACCTCTGGCAAAACGCTACATTTTGGAGTCCAGACGACGACCCCACTGA	834
RESULT 5			
AAI60444			
ID	AAI60444	standard; cDNA; 2152 BP.	
XX	AAI60444;		
XX	AC		
XX	22-OCT-2001	(first entry)	
XX	DT		
XX	DE		
XX	Human polynucleotide SEQ ID NO 4433.		
XX	Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;		
KW	peripheral nervous system; neuropathy; central nervous system; CNS;		
KW	Alzheimer's; Parkinson's disease; Huntington's disease; hemostatic;		
KW	amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemoclastic;		
KW	chemokine; thrombolytic; drug screening; arthritis; inflammation;		
XX	leukemia; ss.		
OS	Homo sapiens.		
XX	WO200153312-A1.		
XX	PN		
XX	26-JUL-2001.		
XX	PD		
XX	26-DEC-2000; 2000WO-US34263.		
XX	21-JAN-2000; 2000US-0488725.		
XX	25-APR-2000; 2000US-0552317.		
PR	PR		

PR	09-JUL-2000;	2000US-0598042.	
PR	19-JUL-2000;	2000US-0620312.	
PR	03-AUG-2000;	2000US-0653450.	
PR	14-SEP-2000;	2000US-0662191.	
PR	19-OCT-2000;	2000US-0693036.	
PR	29-NOV-2000;	2000US-0727344.	
XX			
PA	(HYSE-) HYSEQ INC.		
PI	Tang YF, Liu C, Aundri V, Chen R, Ma Y, Qian XB, Ren F, Wang D;		
PI	Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;		
PI	Zhao QH, Zhou P, Goodrich R, Drmanac RT;		
XX			
DR	WPI: 2001-442253/47.		
DR	P-PSDB; AAM41288.		
PT	Novel nucleic acids and polypeptides, useful for treating disorders		
PT	such as central nervous system injuries -		
PS	Claim 1; SEQ ID NO 4433; 10078pp; English.		
XX			
CC	The invention relates to human nucleic acids (AA157798-AA161369) and		
CC	the encoded polypeptides (AAM48642-AAM42213) with nootropic.		
CC	immunosuppressant and cytostatic activity. The polynucleotides are useful		
CC	in gene therapy. A composition containing a polypeptide or polynucleotide		
CC	of the invention may be used to treat diseases of the peripheral nervous		
CC	system, such as peripheral nervous injuries, peripheral neuropathy and		
CC	localised neuropathies and central nervous system diseases, such as		
CC	Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic		
CC	lateral sclerosis, and Shy-Drager Syndrome. Other uses include the		
CC	utilization of the activities such as: Immune system suppression,		
CC	Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic		
CC	and chromolytic activity, cancer diagnosis and therapy, drug screening,		
CC	assays for receptor activity, arthritis and inflammation, leukaemias and		
CC	C.N.S disorders.		
CC	Note: The sequence data for this patent did not form part of the printed		
CC	specification.		
XX			
XX			
SEQ	Sequence 2152 BP; 468 A; 581 C; 660 G; 442 T; 1 other;		
Query Match	99.8%; Score 832.4; DB 22; Length 2152;		
Best Local Similarity	99.9%; Pred. No. 2.5e-205;		
Matches	833; Conservative 0; Mismatches 1; Indels 0; Gaps 0.		
QY	1 ATGGCTTGGGGGAGGAGGAAGCTGCGAGAGCCCGCGCGGAGAGCGGATCCGAC	60	
DB	59 ATGGCTTGGGGGAGGAGGAGGAAGCTGCGAGAGCCCGCGCGGAGAGCGGATCCGAC	118	
QY	61 CAGCGGCGCTTCTGATAGGGGTGAGCGCGCGGACTGCGAGCGGGAATCGACCGTGT	120	
DB	119 CAGCGGCGCTTCTGATAGGGGTGAGCGCGCGGACTGCGAGCGGGAATCGACCGTGT	178	
QY	121 GAGGAAGATCATGAGATTGCTTGGGACAGAACGAGGTGGAACGCGGACCGGAAGTGGTC	180	
DB	179 GAGGAAGATCATGAGATTGCTTGGGACAGAACGAGGTGGAACGCGGACCGGAAGTGGTC	238	
QY	181 ATCTGAGCGCCAGGACAGGTTTCAACAGGTCCTGACCGGACGAGCAAGAGCCAGGCTTG	240	
DB	239 ATCTGAGCGCCAGGACAGGTTTCAACAGGTCCTGACCGGACGAGCAAGAGCCAGGCTTG	298	
QY	241 AAGGACAGTATCAATTTTGAACATCCAGATCCCTTGTATATGATTGATGATCAGAGACT	300	
DB	299 AAGGACAGTATCAATTTTGAACATCCAGATCCCTTGTATATGATTGATGATCAGAGACT	358	
QY	301 CTGAAGAACATCTGAGAGGCGCAAAACGGTGAAGTGCAGACTATGATTTTGTGACACAC	360	
DB	359 CTGAAGAACATCTGAGAGGCGCAAAACGGTGAAGTGCAGACTATGATTTTGTGACACAC	418	
QY	361 TCAAGGTTTACAGAGACACAGGTGTCTACCGTCGAGCGGTTCTGTTTGAAGGCGATC	420	
DB	419 TCAAGGTTTACAGAGACACAGGTGTCTACCGTCGAGCGGTTCTGTTTGAAGGCGATC	478	
QY	421 TTGGTGTCTTACACCCAGAGATCCGGGACATGTTTCCACTCGCCCTTTGTGTGACACC	480	

```

Db      479 TTGGTGTCTACAGCCAGAGATCCGGAGACATGTTCCACTGCGCTTCTTGAGACACC 538
Qy      481 GACTCCGAGCTCAGGCTGTCTGAAAGATTCTCCGGGAGCGGCGCGAGGGAGGACCTG 540
Db      539 GACTCCGAGCTCAGGCTGTCTGAAAGATTCTCCGGGAGCGGCGCGAGGGAGGACCTG 598
Qy      541 GAGCAGATTCTGACGACGATACCACTTCTGTAAGCGGCGCTTGAAGAGTTCTGCTG 600
Db      599 GAGCAGATTCTGACGACGATACCACTTCTGTAAGCGGCGCTTGAAGAGTTCTGCTG 658
Qy      601 CCGACAAAGAGTATGCGGATGTATCTATCCACGAGAGTGGACAAATATGTTGCCATC 660
Db      659 CCGACAAAGAGTATGCGGATGTATCTATCCACGAGAGTGGACAAATATGTTGCCATC 718
Qy      661 AACCTGATCGTACGACATCCAGACATTTCTGATGATGTCATCTGCAATATGGACCGA 720
Db      719 AACCTGATCGTACGACATCCAGACATTTCTGATGATGTCATCTGCAATATGGACCGA 778
Qy      721 GAGAGGTCATATGGCGGAGCTACAGCGGACCTTTTCTGAGCGCAAGGGACCACTGGG 780
Db      779 GAGAGGTCATATGGCGGAGCTACAGCGGACCTTTTCTGAGCGCAAGGGACCACTGGG 838
Qy      781 ATGCTGACCTCTGGCAACCGTCACTTTGAGAGTCCAGACAGACCCCACTGA 834
Db      839 ATGCTGACCTCTGGCAACCGTCACTTTGAGAGTCCAGACAGACCCCACTGA 892

```

RESULT 6

ABL89762
ID ABL89762 standard; cDNA; 1396 BP.

XX ABL89762;

DT 24-MAY-2002 (first entry)

DE Human polynucleotide SEQ ID NO 324.

XX Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
 KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor;
 KW antineoplastic; anticonvulsant; antibacterial; antifungal; antiparasitic;
 KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; human; secreted protein; gene; ss.

XX Homo sapiens.

XX OS WO200190304-A2.

XX PN 29-NOV-2001.

XX PF 18-MAY-2001; 2001WO-US16450.

XX PR 19-MAY-2000; 2000US-205515P.

XX PA (HUMA-) HUMAN GENOME SCI INC.

XX PI Birse CE, Rosen CA;

XX DR WPI; 2002-122018/16.

XX DR P-PSDB: ABB89353.

XX Novel 1405 isolated polypeptides, useful for diagnosis, treatment and
 PT prevention of neural, immune system, muscular, reproductive,
 PT gastrointestinal, pulmonary, cardiovascular, renal and proliferative
 PT disorders -

PS Claim 4; SEQ ID NO 324; 2081bp + Sequence Listing; English.

XX The invention relates to novel genes (ABL89449-ABL90853) and proteins
 CC (ABB89040-ABB90444) useful for preventing, treating or ameliorating
 CC medical conditions e.g. by protein or gene therapy. The genes are
 CC isolated from a range of human tissues disclosed in the specification.
 CC The nucleic acids, proteins, antibodies and (ant)agonists are useful

CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
 CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
 CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
 CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
 CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
 CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
 CC colitis; (c) cardiovascular disorders such as myocardial ischaemia;
 CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
 CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
 CC and parasitic infections.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 1396 BP; 317 A; 358 C; 425 G; 287 T; 9 other;

Query Match 97.0%; Score 808.8; DB 24; Length 1396;
 Best Local Similarity 99.4%; Pred. No. 2.8e-199;
 Matches 829; Conservative 3; Mismatches 0; Indels 2; Gaps 2;

```

Qy      1 ATGGCTTCGCGGAGGCGAAGACTGCGAGAGCCCGCGCGGAGCCGACCTCCGAC 60
Db      41 ATGGCTTCGCGGAGGCGAAGACTGCGAGAGCCCGCGCGGAGCCGACCTCCGAC 100
Qy      61 CAGCGGCTTCTCTGATAGGGGTGAGCGGCGCACTGCCAGCGGAAATGACCTGTGT 120
Db      101 CAGCGGCTTCTCTGATAGGGGTGAGCGGCGCACTGCCAGCGGAAATGACCTGTGT 160
Qy      121 GAGAAATCATATGAGTTCGTCGAGACAGAACAGAGTGAACAGCGGACCGGAGTGTCT 180
Db      161 GAGAAATCATATGAGTTCGTCGAGACAGAACAGAGTGAACAGCGGACCGGAGTGTCT 220
Qy      181 ATCCTGAGCCAGAGACAGTTCCTCAAGGTCCTACGCGCAGAGCAGAGCCAGGCTTG 240
Db      221 ATCCTGAGCCAGAGACAGTTCCTCAAGGTCCTACGCGCAGAGCAGAGCCAGGCTTG 280
Qy      241 AAGGACATGATATTTGACATCCAGATGCTTTGATATGATTTGATGACAGAGCT 300
Db      281 AAGGACATGATATTTGACATCCAGATGCTTTGATATGATTTGATGACAGAGCT 340
Qy      301 CTGAAGAACTCGTGAAGGGCAAAACGGTGAAGTCCCACTATATTTTGGACACAC 360
Db      341 CTGAAGAACTCGTGAAGGGCAAAACGGTGAAGTCCCACTATATTTTGGACACAC 400
Qy      361 TCAAGTTACAGAGACCAAGGTGTCTACCTCGCGAGCGTGTCTGTTTGAAGGCATC 420
Db      401 TCAAGTTACAGAGACCAAGGTGTCTACCTCGCGAGCGTGTCTGTTTGAAGGCATC 460
Qy      421 TTGGTGTCTACAGCAGCAGATCCGAGACATGTCACCTGCGCTCTTCTGTGACACC 480
Db      461 TTGGTGTCTACAGCAGCAGATCCGAGACATGTCACCTGCGCTCTTCTGTGACACC 520
Qy      481 GACTCCGAGCGTCAAGGCTGTCTGAAGATCTCCGAGAGTGGCCGAGGAGGAGACTG 540
Db      521 GACTCCGAGCGTCAAGGCTGTCTGAAGATCTCCGAGAGTGGCCGAGGAGGAGACTG 580
Qy      541 GAGCAGATTCTGACGAGTACACACCTTCTGTAAGCCCGGCTTTCGAGAGATTCTGCTG 600
Db      581 GAGCAGATTCTGACGAGTACACACCTTCTGTAAGCCCGGCTTTCGAGAGATTCTGCTG 639
Qy      601 CCGACAAAGAGTATCCGATGTATCCACGAGAGTGAACAATATGTTGCCATC 660
Db      640 CCGACAAAGAGTATCCGATGTATCCACGAGAGTGAACAATATGTTGCCATC 699
Qy      661 AACCTGATGTCAGACATCCAGACATTTCTGAATGTCATCTGCAATATGTCACCGA 720
Db      700 AACCTGATGTCAGACATCCAGACATTTCTGAATGTCATCTGCAATATGTCACCGA 759
Qy      721 GAGAGGTCATATGGCGGAGCTACAGCGGACCTTTTCTGAGCGAGGAGCACCCTGGG 780
Db      760 GAGAGGTCATATGGCGGAG-TCACAGCGGACCTTTTCTGAGCGAGGAGCACCCTGGG 818
Qy      781 ATGCTGACCTCTGGCAACCGTCACTTTGAGAGTCCAGACAGACCCCACTGA 834

```

Db 819 ATGCTGACCTCTGCAAAACGTCATTTGGAGTCCAGACGACGACCCCACTGA 872

|||||

RESULT 7
AAH75355
AAH75355 standard; cDNA; 900 BP.

XX AC AAH75355;
XX 02-OCT-2001 (first entry)
XX Human uridine kinase encoding cDNA.
XX Human; uridine kinase; UK; ss.
XX Homo sapiens.
XX OS
XX Key Location/Qualifiers
XX CDS 8..838
XX FT /*tag= a
XX FT /product= "uridine kinase"
XX FT /note= "Claimed in claim 1"
XX
XX CN1287172-A.
XX 14-MAR-2001.
XX 07-SEP-1999; 99CN-0118818.
XX 07-SEP-1999; 99CN-0118818.
XX (UYFU-) UNIV FUDAN.
XX Yu L, Zhao Y, Zhang H;
XX WPI; 2001-409529/44.
XX P-PSDB; AAG64506.
XX
XX Human uridine kinase and its coding sequence, preparation and
XX application -
XX
XX Claim 1; Page 15 (Disclosure); 20pp; Chinese.
XX
XX The invention relates to human uridine kinase (UK).
XX
XX Sequence 900 BP; 201 A; 237 C; 292 G; 170 T; 0 other;
XX
XX Query Match 94.2%; Score 786; DB 22; Length 900;
XX Best Local Similarity 97.2%; Pred. No. 1.9e-193;
XX Matches 811; Conservative 0; Mismatches 20; Indels 3; Gaps 1;

QY 1 ATGGCTTCGGCGGAGGCGAAGACTGCGAGACCCCGCGGAGCGGACCGTCCGAC 60
DB 8 ATGGCTTCGGCGGAGGCGAAGACTGCGAGACCCCGCGGAGCGGACCGTCCGAC 64

QY 61 CAGCGGCGCTTCTCTGATAGGGGTGAGCGCGGCACTCCAGCGGAAAGTGCACCTGTGT 120
DB 65 CACGCTCCCTTCTCTGATAGGGGTGAGCGCGGCACTCCAGCGGAAAGTGCACCTGTGT 124

QY 121 GAGAGATCATGAGTGTCTGGACAGAAAGAGGTGGAAGAGCGGCGCGGAAGGTGCTC 180
DB 125 GAGAGATCATGAGTGTCTGGACAGAAAGAGGTGGAAGAGCGGCGCGTAAGGTGCTC 184

QY 181 ATCTGAGCCAGGACAGTTCCTACAGGCTCTGACGCGACAGGACGAGGCGCTTG 240
DB 185 ATCTGAGCCAGGACAGTTCCTACAGGCTCTGACGCGACAGGACGAGGCGCTTG 244

QY 241 AAGAGACATGATTTTGAACATCCAGATCCCTTTGATATGATTTGATGACAGGACT 300
DB 245 AAGAGACATGATTTTGAACATCCAGATCCCTTTGATATGATTTGATGACAGGACT 304

QY 301 CTGAAGAACTGTCGAGGCGCAAAACGCTGAGAGTGCAGACTATGATTTTGTGACACAC 360

Db 305 CTGAAGAACTGTCGAGGCGCAAAACGCTGAGAGTCCCGACCTATGATTTTGTGACACAC 364

QY 361 TCAAGTTACAGAGACCAAGGTGTCTACCTTCGCGAGAGTGTCTGTTTGAAGGCATC 420
DB 365 TCAAGTTACAGAGACCAAGGTGTCTACCTTCGCGAGAGTGTCTGTTTGAAGGCATC 424

QY 421 TTGGTGTCTACAGCAGGAGATCCGGGACATGTTCCACTGGCCCTTCGTGGAACAC 480
DB 425 TTGGTGTCTACAGCAGGAGATCCGGGACATGTTCCACTGGCCCTTCGTGGAACAC 484

QY 481 GACTCCGACGTCAAGCTGTCTCGAAGATTTCTCCGGAAGTGGCCGAGAGGAGACCTG 540
DB 485 GACTCCGACGTCAAGCTGTCTCGAAGATTTCTCCGGAAGTGGCCGAGAGGAGACCTG 544

QY 541 GAGCAGATTTCTGACGACATCACACCTTCGTGAGACCGGCTTCGAGAGATTTCTGCTG 600
DB 545 GAGCAGATTTCTGACGACATCACACCTTCGTGAGACCGGCTTCGAGAGATTTCTGCTG 604

QY 601 CCGACAAAGATATCCGATGATCATCCGACGAGAGTGGACAAATATGTTGCAATC 660
DB 605 CCGACAAAGATATCCGATGATCATCCGACGAGAGTGGACAAATATGTTGCAATC 664

QY 661 AACCTGATGTCGACGACATCCAGACATTTGATGATGTCATGCAAAATGGACCGA 720
DB 665 AACCTGATGTCGACGACATCCAGACATTTGATGATGTCATGCAAAATGGACCGA 724

QY 721 GAGGATCCAAATGGCGGAGCTCAAGCGGACCTTTTCTGAGCCAGGAGACCACTGGG 780
DB 725 GAGGATCCAAATGGCGGAGCTCAAGCGGACCTTTTCTGAGCCAGGAGACCACTGGG 784

QY 781 ATGCTGACCTCTGCGCAACGCTCACTTGGAGTCCAGAGACGACCCCACTGA 834
DB 785 ATGCTGACCTCTGCGCAACGCTCACTTGGAGTCCAGAGACGACCCCACTGA 838

RESULT 8
AAH04832
AAH04832 standard; cDNA; 753 BP.

XX AC AAH04832;
XX
XX 26-JUN-2001 (first entry)
XX
XX Human cDNA clone (5'-primer) SEQ ID NO:1667.
XX
XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
XX Homo sapiens.
XX
XX EPI074617-A2.
XX
XX 07-FEB-2001.
XX
XX 28-JUL-2000; 2000BP-0116126.
XX
XX 29-JUL-1999; 99JP-0248036.
XX 27-AUG-1999; 99JP-0300253.
XX 11-JAN-2000; 2000JP-0118776.
XX 02-MAY-2000; 2000JP-0183767.
XX 09-JUN-2000; 2000JP-0241899.
XX
XX (HELI-) HELIX RES INST.
XX
XX Ota T, Isegai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
XX Iehi S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX WPI; 2001-318749/34.
XX
XX Primer sets for synthesizing polynucleotides, particularly the 5602
XX full-length cDNAs defined in the specification, and for the detection
XX and/or diagnosis of the abnormality of the proteins encoded by the
XX full-length cDNAs -

XX Claim 1; SEQ ID 1667; 2537bp + CD ROW; English.
XX
XX The present invention describes primer sets for synthesizing 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.
XX
XX Sequence 753 BP; 158 A; 196 C; 261 G; 135 T; 3 other;
SQ
Query Match 74.8%; Score 623.6; DB 22; Length 753;
Best Local Similarity 97.7%; Pred. No. 1.9e-151;
Matches 642; Conservative 0; Mismatches 12; Indels 3; Gaps 1;
QY 1 ATGGCTTCGGGGGAGGCGAAGACGAGAGCCCGGCGGAGGCGGACCGCCGAC 60
Db 95 ATGGCTTCGGGGGAGGCGAAGACGAGAGCCCGGCGGAGGCGGACCGCCGAC 154
QY 61 CAGCGGCCCTTCTGATAGGGGTGAGCGCGCACTGCCAGGGAAAGTCAGCCGTGT 120
Db 155 CAGCGGCCCTTCTGATAGGGGTGAGCGCGCACTGCCAGGGAAAGTCAGCCGTGT 214
QY 121 GAGAAATCATGAGTTCTCTGGACAAACAGAGTGGAAACGCGGACGGAAGGTGTC 180
Db 215 GAGAAATCATGAGTTCTCTGGACAAACAGAGTGGAAACGCGGACGGAAGGTGTC 274
QY 181 ATCTGAGCCGAGCAGAGTCTCAACAGGTCTCTGACGAGAGCAGAAAGGCCATCTG 240
Db 275 ATCTGAGCCGAGCAGAGTCTCAACAGGTCTCTGACGAGAGCAGAAAGGCCATCTG 334
QY 241 AAAGGACAGTCAATTTTGAACCATCCAGATGCTTTGATAATGATTGATGACAGACT 300
Db 335 AAAGGACAGTCAATTTTGAACCATCCAGATGCTTTGATAATGATTGATGACAGACT 394
QY 301 CTGAAGAATCTGCGAGGGGCAAAACGTTGAGGTGCCGACTTATGATTGTGACACAC 360
Db 395 CTGAAGAATCTGCGAGGGGCAAAACGTTGAGGTGCCGACTTATGATTGTGACACAC 454
QY 361 TCAAGTTTACAGAGACCAAGGATGCTCAACCTCGGACGAGTTCGTTTGAAGGGATC 420
Db 455 TCAAGTTTACAGAGACCAAGGATGCTCAACCTCGGACGAGTTCGTTTGAAGGGATC 514
QY 421 TTGGTGTCTACACAGCAGAGATCCGGGACATGTTTCCACCTGCGCTTCTGTGACACC 480
Db 515 TTGGTGTCTACACAGCAGAGATCCGGGACATGTTTCCACCTGCGCTTCTGTGACACC 574
QY 481 GACTCCGACGTCAGGCTGTCTCGAAGAGTTCTCCGGACGTGCGCCGAGGGAGGACCTG 540
Db 575 GACTCCGACGTCAGGCTGTCTCGAAGAGTTCTCCGGACGTGCGCCGAGGGAGGACCTG 634
QY 541 GAGCAGATTCTGAGCAGTACACACACCTTGTTGAAGCGCGCTTCGAGAGTTCTGACC 598
Db 635 GAGCAGATTCTGAGCAGTACACACACCTTGTTGAAGCGCGCTTCGAGAGTTCTGACC 694
QY 599 -TGCAGCAAAAGAGTATGCCGATGTGATCATCCACAGAGAGTGAACAATATGTT 654

Db 695 TGNCGACAAAGAGTGTGCGGATGATCATTCACAGAGAGTGAACAATATGTT 751
|||||
RESULT 9
ABBS1801
ID ABBS1801 standard; cDNA; 734 BP.
XX
AC ABBS1801;
XX
DT 21-OCT-2002 (first entry)
XX
DE Human mddt cDNA Incyte ID No: LI:235557.12:2001JAN12.
XX
KW Human; molecule for disease detection and treatment; MDDT; cancer;
KW cell proliferative disorder; arteriosclerosis; cirrhosis; asthma;
KW autoimmune disorder; inflammatory disorder; Crohn's disease;
KW multiple sclerosis; cytostatic; antiarteriosclerotic; antiinflammatory;
KW hepatotropic; immunosuppressive; antiaesthetic; gene therapy; gene; ss.
XX
OS Homo sapiens.
XX
PN W020025738-A2.
XX
PD 18-JUL-2002.
XX
PF 09-JAN-2002; 2002MO-US01008.
XX
PR 12-JAN-2001; 2001US-261622P.
PR 16-JAN-2001; 2001US-261865P.
PR 17-JAN-2001; 2001US-262208P.
PR 17-JAN-2001; 2001US-262209P.
PR 17-JAN-2001; 2001US-262326P.
PR 19-JAN-2001; 2001US-263065P.
PR 19-JAN-2001; 2001US-263065P.
PR 19-JAN-2001; 2001US-263329P.
XX
PA (INCYTE) INCYTE GENOMICS INC.
XX
PI Panzer SR, Lincoln SE, Altus CM, Dufour GE, Hillman JL, Jones AL;
PI Dam TC, Liu TF, Harris B, Flores V, Dafo A, Marwaha R, Chen AD;
PI Chang SC, Gerstlin EH, Peralta CH, David MH, Lewis SA;
XX
DR WPI; 2002-590679/63.
DR P-PADB; ABG70328.
XX
PT New disease detection and treatment molecule (MDDT) polynucleotides and
PT polypeptides, useful in diagnosing, studying, preventing or treating
PT diseases associated with MDDT expression, e.g. autoimmune or
PT inflammatory disorders -
XX
PS Claim 1; Page 98; 129pp; English.
XX
XX The present invention relates to the isolation of novel human
CC molecules for disease detection and treatment (MDDT), and the
CC polynucleotide sequences (mddt) encoding them. The MDDT polypeptides
CC may be used to screen for molecules that bind to, or are bound by the
CC encoded polypeptides, and to develop a transcript image of a tissue or
CC cell type. Probes comprising at least 20 nucleotides of the mddt
CC polynucleotide may be used to assess the toxicity of a test compound.
CC The MDDT polypeptides and mddt polynucleotides are useful in the
CC diagnosis, study, prevention and treatment of diseases associated with
CC the expression of molecules for disease detection and treatment. Such
CC disorders include cell proliferative disorders (e.g. arteriosclerosis,
CC cirrhosis, or cancers), and autoimmune/inflammatory disorders
CC (e.g. asthma, Crohn's disease, or multiple sclerosis). The mddt
CC polynucleotides may also be used as molecule markers, in microarrays,
CC and in somatic or germline gene therapy. ABBS1779-ABBS1814 encode
CC the MDDT proteins of the invention.
XX
SQ Sequence 734 BP; 166 A; 181 C; 226 G; 154 T; 7 other;
Query Match 58.0%; Score 483.6; DB 24; Length 734;

Db 426 GAGCCAGATAGCTTCTACCGTGCTTACTCTCGAGCAGAGGCCAAAGCCGTAAAGG 485
Qy 246 ACAGTACAAATTTTACCCATCCAGATGCTTTGATTAATTTGATGACAGAGCTTGAA 305
Db 486 CCAGTTCAACTTTGACACCCGAGTCCCTTGAACAATTAATCTATTCAAAACACTCA 545
Qy 306 GAACATCGTGAAGGCAAAACGCTGAGAGTCCGACCTATGATTTTGTGACACACTCAAG 365
Db 546 AGAATCATGAAGGAAACAGTCCAGATCCCGTGATATACCTTTGTCTCCATCCCG 605
Qy 366 GTTACCAAGACCAACGCTGTCTTACCTCGAGAGCTGTTTGTGAGGCACTTGCT 425
Db 606 GAAGAGAGAGACAGTTACTGTCTTATCCCGCAGACGCTGTGCTTTGAGAGGATCTGGC 665
Qy 426 GTTCTACAGCCAGAGAGATCCGGGACATGTTCACCTGGCCTCTTGGGACACCGATC 485
Db 666 CTTTCTACTCCAGAGAGTACAGACCTTTCAGATGAAGCTTTTGTGATACAGATGC 725
Qy 486 CGACGTCAAGCTGTCTCGAAGATTTCTCCGGGACGTCGC--CGAGGAGAGACCTGGA 542
Db 726 GGAACACCCGGCTCTCAGCAGAGATTAAGGACATCGCAGAGAGGACAGGATCTTGA 785
Qy 543 GCAGATTCTGACGAGTACACCACTTGTGAAAGCCGCTTCAGAGATTTCTGCTGCC 602
Db 786 GCAGATTTTATCTCAGTACATTAAGTTCGTCAAGCTGCTTTGAGGATTTCTGCTGCC 845
Qy 603 GACAAAGAGTATGCCGATGTGATCATCCGACGAGAGTGAACAATATGTTGCACTCA 662
Db 846 AACAAAGAGTATGCTGATGTGATCATCTTGAAGGTGCAGATATCTGGGCACTCA 905
Qy 663 CCTGATCGTGCAGACATCCAGACATTTCTGAATGTGACATCTGCAATGCA 716
Db 906 CCTCATCGTGCAGACATCCAGACATCTGATGAGAGGCGCTTCAAAACGCA 959

RESULT 11

AA157850 standard; cDNA; 1402 BP.

AA157850;

22-OCT-2001 (first entry)

Human polynucleotide seq ID NO 53.

Human: nootropic; immunosuppressant; cyrostatic; gene therapy; cancer;
peripheral nervous system; neuropathy; central nervous system; CNS;
Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
leukaemia; ss.

Homo sapiens.

WO20015312-A1.

26-JUL-2001.

26-DEC-2000; 2000WO-US34263.

21-JAN-2000; 2000US-0488725.
25-APR-2000; 2000US-0552317.
09-JUL-2000; 2000US-0598042.
19-JUL-2000; 2000US-0620312.
03-AUG-2000; 2000US-0653450.
14-SEP-2000; 2000US-0662191.
19-OCT-2000; 2000US-0693036.
29-NOV-2000; 2000US-0727344.

(HYSE-) HYSEQ INC.

Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;

PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX WPI; 2001-442253/47.
DR P-PSDB; AAM38694.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
XX
PS Claim 1; SEQ ID NO 53; 10078pp; English.

XX The invention relates to human nucleic acids (AA157798-AA161369) and
CC the encoded polypeptides (AAM38642-AA42213) with nootropic,
CC immunosuppressant and cyrostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemia and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.

SQ Sequence 1402 BP; 379 A; 370 C; 371 G; 282 T; 0 other;

Query Match 44.8%; Score 374; DB 22; Length 1402;

Best Local Similarity 74.3%; Pred. No. 8.8e-87;

Matches 486; Conservative 0; Mismatches 165; Indels 3; Gaps 1;

Qy 66 GCCCTTCTGATAGGGGTGAGCCGCCGACCTGCCAGCGGAATCGACCTGTGTGAA 125
Db 346 GCCCTTCTTATAGGGGTGAGCCGCCGGAACAGCTACCGCACTCTTCCGTGTCTAA 405
Qy 126 GATCATGAGATGCTGTGGACAGAACGAGGTGAACGCGGACGCGAAGGTGTCATCCT 185
Db 406 GATGTGACAGCTCTGTGGGACGAATAGGTGACTTTCGCAAGACAGGTGTCATCCT 465
Qy 186 GAGCCAGACAGGTTCTACAAAGTCTCTGACGCGACAGACAAAGCCCTTGAAG 245
Db 466 GAGCCAGATAGCTTACCGGTCTTACCTCTGAGGACGAAGGCCAAAGCCCTGAAG 525
Qy 246 ACAGTACATTTTACCATTCAGATGCTTTGATTAATGATGATGACAGACTTGAA 305
Db 526 CCAGTTCAACTTTGACACCCGAGATCCTTTGACAAATGAACTCATCTCAAAACACTCA 585
Qy 306 GAACATCGTGAAGGCAAAACGCTGAGAGGCCGACCTTATGATTTTGTGACACACTCA 365
Db 586 AGAATCATGAAGGAAACAGTCCAGATCCCGTGATGACTTTGTCTCCATTCGG 645
Qy 366 GTTACCAAGACCAACGATGCTTACCTCGGACGATGTTCTGTTGAGGCACTTGGT 425
Db 646 GAAGAGAGAGACGTTACTGTCTATCCGGAACGATGCTTTGAAAGGATCTTGGC 705
Qy 426 GTTCTACAGCCAGAGATCCGGAACATGTTCCACCTGCGCTTCTTGTGACACGACTC 485
Db 706 CTTTCTACTCCAGAGAGTACGAACTGTTCAGATGAAGCTTTTGTGATTAACATATC 765
Qy 486 CGAGTCAAGCTGTCTCGAAGTTCCTCGGGACGTCGC---CGAGGAGGAGACTTGA 542
Db 766 GGAACACCCGGCTTCCAGAGATATTAAGGACATCAACGAGAGGACGAGATTTGA 825
Qy 543 GCAGATTCTGACGACATACCACTTCTGTGAAGCCGCTTCAGAGATTTCTGCTGCC 602
Db 826 GCAGATTTTATCATGATACATTAAGTTCGCAAGCTGCTTTGAGGAATTCGCTGCC 885
Qy 603 GACAAAGAGTATGCCAGTGTGATCATCCACGAGAGTGAACATATGTTGCCATCA 662
Db 886 AACAAAGAGTATGCTGATGTGATCATCTTGAAGGTGACAGATTAATCTGTGCTCA 945

Qy 663 CCTGATGTCAGACATCCAGACATTTGATGTGACATCTGCAATGCGA 716
 Db 946 CCTGATGTCAGACATCCAGACATCTGATGATGAGGGCCCTCCAAACGCA 999

RESULT 12

AAF15785
 ID AAF15785 standard; cDNA; 1310 BP.

AAF15785;

13-MAR-2001 (first entry)

Human prostate cancer antigen nucleotide sequence SEQ ID NO:220.

Human prostate cancer; prostate cancer antigen; detection; diagnosis;
 neuroprotective; cytoskeletal; cardioactive; immunomodulatory; muscular;
 vulnery; gastrointestinal; nephrotoxic; antiinfective; gynaecological;
 antimicrobial; gene therapy; neural; immune; reproductive; renal;
 gastrointestinal; pulmonary; cardiovascular; proliferative disorder;
 wound; infectious disease; ss.

Homo sapiens.

MO200055174-A1.

21-SEP-2000.

08-MAR-2000; 2000WO-US05988.

12-MAR-1999; 99US-0124270.

(HUMA-) HUMAN GENOME SCI INC.

(ROSE/) ROSEN C A.

Rosen CA, Ruben SM;

WPI; 2000-587513/55.

P-PSDB; AAB56582.

Prostate cancer associated gene sequences, referred to as prostate
 cancer antigens, useful for treatment, prevention, and diagnosis of
 disorders such as prostate cancer -

Claim 1; Page 752; 2338bp; English.

AAF15566 to AAF16505 encode the human prostate cancer associated

proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.

The prostate cancer antigens can have neuroprotective, cytoskeletal,

cardioactive, immunomodulatory, muscular, vulnery, gastrointestinal,

nephrotoxic, antiinfective, gynaecological and antibacterial activities,

and can be used in gene therapy. The prostate cancer antigen

polynucleotides may be used for detection of prostate cancer, chromosome

identification, as chromosome markers, and for numerous other diagnostic

or research purposes. The prostate cancer antigens may be used to treat

disorders such as neural, immune, muscular, reproductive,

gastrointestinal, pulmonary, cardiovascular, renal, and proliferative

disorders, wounds, and infectious diseases. AAF15566 to AAF16514 to

AAF57303 represent sequences used in the exemplification of the present

invention.

Sequence 1310 BP; 338 A; 350 C; 352 G; 264 T; 6 other;

Query Match 43.4%; Score 361.6; DB 21; Length 1310;

Best Local Similarity 74.0%; Pred. No. 1,4e-83;

Matches 484; Conservative 1; Mismatches 165; Indels 4; Gaps 2;

66 GCCCTTCCTGATGAGGCGGCGGCACTGCGAGGGAAGTCCAGCGTGTGAGAA 125

288 GCCCTTCCTGATGAGGCGTCA-CGGGGAACAGCTAGGCGCAAGCTTCCGTGTGTAA 346

126 GATCATGAGATTCTGGGACAGACGAGTGAACAGCGGACGGAAGGTGTATCCT 185

Db 347 GATCTGTCAGACTCCTGGGGGAGATGAGTGAAGTATCCGACAGAGAGTGTATCT 406
 Qy 186 GAGCCAGACAGGTTCTTCAAGAGTCTGACGGGAGAGACAGAAAGCCCTTGAAGG 245
 Db 407 GAGCCAGAGATGCTTCAAGGTCCTTCACTCTGACCTGAGGAGAAAGCCCTGAAGG 466
 Qy 246 ACAGTCAATTTTGAACATCCAGATCCCTTGAATTAATGATTTGACAGACTCAAG 305
 Db 467 CCAATTCACTTGAACACCCGAGATCCCTTGAATTAATGATTTGACAGACTCAAG 526
 Qy 306 GAAATCACTGAGAGGAGAAACAGTGAAGTCCGACCTATGATTTTGTGACACTCAAG 365
 Db 527 GAAATCACTGAGAGGAGAAACAGTGAAGTCCGACCTATGATTTTGTGACACTCAAG 586
 Qy 366 GTTACCAAGACCAAGTGTCTTCACTCTGACCTGAGGATTTTGTGACACTCAAG 425
 Db 587 GAAAGAGAGACAGTATCTGATCTTCACTCTGACCTGAGGATTTTGTGACACTCAAG 646
 Qy 426 GTTCTCAAGCAGAGATCCGAGACATGTTCCACCTGCGCTCTTGTGAGACCCGACTC 485
 Db 647 CTCTTACTCCAGAGAGTACAGACTGTTCCAGATGAAGCTTTTGTGATACAGATGC 706
 Qy 486 CGACGTCAAGCTGTCTGAAAGATTTCCGAGACGTGCGC---CGAGGAGGAGCTGGA 542
 Db 707 GACACCCGCTCTCAAGAGATTTTAAAGGACATCAGAGAGAGGAGGAGATCTTGA 766
 Qy 543 GCGATTTCTGACGAGTACCACTTCTGAAAGCCGCTTGTGAGAGTCTTGTGCTGCC 602
 Db 767 GCGATTTTATCTCAAGTACATTTACGTTGCAAGCTGCTTGTGAGAGATTTGTGCTGCC 826
 Qy 603 GACAAAGATATGCGATGATGATCCACAGAGAGTGAACAATATGTTGCCATCA 662
 Db 827 AACAAAGATATGCTGATGATGATCTTCTGAGAGTGAACAATATGTTGCCATCA 886
 Qy 663 CCTGATGTCAGACATCCAGACATTTGATGTGACATCTGCAATGCGA 716
 Db 887 CCTGATGTCAGACATCCAGACATCTGATGATGAGGGCCCTCCAAACGCA 940

RESULT 13

ABQ54470
 ID ABQ54470 standard; cDNA; 1322 BP.

ABQ54470;

22-AUG-2002 (first entry)

Human ovarian antigen HOBLP29 cDNA, SEQ ID NO:350.

Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;
 ovarian cancer; breast cancer; tumour; reproductive system disorder;
 infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
 PCOS; ovarian cyst; dysmenorrhoea; endocrine disorder; infection;
 inflammatory condition; immune disorder; blood disorder;
 cardiovascular disorder; respiratory disorder; neurological disorder;
 gastrointestinal disorder; urinary system disorder; drug screening;
 gene therapy; chromosome mapping; forensic analysis;
 antibody preparation; cytoskeletal; immunomodulatory; neuroprotective;
 antiinflammatory; gynaecological; reproductive; chromosome 1p32;
 gene; ss.

Homo sapiens.

WO200200677-A1.

03-JAN-2002.

07-JUN-2001; 2001WO-US16569.

07-JUN-2000; 2000US-209467P.

(HUMA-) HUMAN GENOME SCI INC.

PI Birse CE, Rosen CA;
 XX
 XX WPI: 2002-147878/19.
 DR P-PSDB; ABP41393.
 XX
 PT Isolated nucleic acid molecules encoding novel ovarian polypeptides,
 PT useful in the prevention, treatment and diagnosis of cancer (e.g.
 PT ovarian cancer), immune disorders, cardiovascular disorders and
 PT neurological diseases -
 XX
 PS Claim 1; SEQ ID NO 350; 29222bp; English.

XX
 XX The invention relates to 2175 novel human ovarian antigens (ABP41054-
 CC ABP43228) and to cDNAs encoding them (ABO54131-ABO56305), and also
 CC encompasses polypeptides 90% identical and polynucleotides 95% identical
 CC to the sequences of the invention. The invention additionally relates to
 CC recombinant vectors and host cells comprising human ovarian antigen
 CC polynucleotides, antibodies against human ovarian antigens, and the use
 CC of ovarian antigen polynucleotides and polypeptides in diagnosing,
 CC treating, prognosing or preventing various ovarian and/or breast-related
 CC disorders. Such conditions include ovarian cancer and breast cancer, and
 CC metastatic tumours of ovarian or breast origin, reproductive system
 CC disorders (e.g., infertility, disorders of pregnancy, anovulation,
 CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine
 CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic
 CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
 CC vaginitis), immune disorders (e.g., congenital and acquired
 CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
 CC blood-related disorders (e.g., anaemia), cardiovascular disorders,
 CC respiratory disorders, neurological disorders, gastrointestinal disorders
 CC and urinary system disorders. Ovarian antigen polypeptides and
 CC polynucleotides may also be used in screening for compounds which
 CC modulate ovarian antigen expression or activity. The polynucleotides may
 CC further be used for gene therapy, chromosome mapping, in the
 CC identification of individuals and in forensic analysis, and the
 CC polypeptides may be used as food additives or to prepare antibodies
 CC useful in disease diagnosis, drug targeting and phenotyping. The present
 CC sequence represents cDNA encoding a human ovarian antigen of the
 CC invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC
 XX

SQ Sequence 1322 BP; 351 A; 351 C; 352 G; 264 T; 4 other;

Query Match 43.4%; Score 361.6; DB 24; Length 1322;
 Best Local Similarity 74.0%; Pred. No. 1.4e-83;
 Matches 484; Conservative 1; Mismatches 165; Indels 4; Gaps 2;

QY 66 GCCCTTCTGATAGGGGTGAGCGGCGGACCTGCCAGCGGGAAGTCACCGTGTGAGAA 125
 |||||
 DB 288 GCCCTTCTTATAGGGCGTCA-CGGGGGAACAGCTAGCGGAGCTTCCGTGTGCTAA 346
 QY 126 GATATGAGATTGCTGTGGACGAACGAGGTGAAACAGCGGACGGGAGGTGATCTT 185
 |||||
 DB 347 GATGTGACAGTCTCTGGGGGAGATGAGTGAATCTATGCGCAGAGGAGGTGATCTT 406
 QY 186 GAGCCAGGACAGGTTCTTACAGGTCTGACGCGACAGACGAGGCGCAAGCGCTTGAAGG 245
 |||||
 DB 407 GAGCGAGGATAGCTTCTTACCGCTGTCTTACCTCGAGAGGAGGCGCAAGCGCTGAAGG 466
 QY 246 ACAATGAAATTTTACCATCCAGATGCTTGTATGATGATTGATGACAGACTTGA 305
 |||||
 DB 467 CCAATTCACATTTTACCAACCGGATGCTTGTGACATGAATCACTTTCTCAAAACACTCAA 526
 QY 306 GAACATGCTGAGGGCAAAACGCTGAGGTGCGGACCTGATGTTTGTGACACACTCAAG 365
 |||||
 DB 527 AGAATACCTGAAAGGAAACAGTCCAGATCCCGTGATGACTTTGTCTCCCATCTCCG 586
 QY 366 GTTACAGAGACCAACGCTGCTACCTCGGAGAGTGTGTTGAGGGCATCTTG 425
 |||||
 DB 587 GAAGGAGAGACAGTACTGCTTATCTCCGACAGAGTGTGTCTTTGAAGGAGCTCTGCG 646

QY 426 GTTCTACAGCAGAGATCCGGACATGTTCCACTCGGCTCTGCTGAGACCCGACTC 485
 |||||
 DB 647 CTCTTACTCCAGAGAGGTACGAGACCTGTTCCAGATGAACTTTTGTGATACAGATGC 706
 QY 486 CAGAGTCAGGCTGTCTTGAAGAGTTCTCCGGACGTGCGC--CGAGGAGGACCTTGA 542
 |||||
 DB 707 GGAACCCGGCTCTCAGCGAGAGTATTAAGGACATCAGCAGAGAGGACGAGGATCTTGA 766
 QY 543 GCAGTCTTCAACGACATACCAACCTTCGTGAAGCCGCTTGAAGAGATCTTCCCTGCC 602
 |||||
 DB 767 GCAGATTTTATTCAGTACATTAAGTTGTTGCAAGCCCTGTTGAGGAAATCTGCTTGGC 826
 QY 603 GACAAAGAAATAGCCGATGTGATCATCCACGAGAGTGACAAATATGTTGCCATCAA 662
 |||||
 DB 827 MACAAAGAAATAGCTGATGTGATCATCCCTTAAGGTGACGATTAATCTGTTGCCATCAA 886
 QY 663 CTTGATCTGACAGACATCCAGACATTTCTGAATGTTGACATCTGCAATGGA 716
 |||||
 DB 887 CTTGATCTGACAGACATCCAGACATCTGAAATGAGGGCCCTCCAAACGGCA 940

RESULT 14

ABLI0981
 ID ABLI0981 standard; cDNA; 822 BP.
 XX
 AC ABLI0981;
 XX
 DT 26-MAR-2002 (first entry)
 DE
 XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 27425.
 DE
 XX Drosophila; developmental biology; cell signalling; insecticide;
 XX pharmaceutical; gene; ss.
 XX
 KW Drosophila melanogaster.
 OS
 XX
 PN WO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US09231.
 XX
 PR 23-MAR-2000; 2000US-191637P.
 XX
 PR 11-JUL-2000; 2000US-0614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PWD, Myers EW;
 XX
 DR WPI: 2001-656860/75.
 DR P-PSDB; ABB66878.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 PT
 PS Claim 1; SEQ ID NO 27425; 212p + Sequence Listing; English.
 XX
 XX The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABLI16176-ABLI30511), expressed DNA
 CC sequences (ABLI1840-ABLI16175) and the encoded proteins
 CC (ABBS7737-ABBS7072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC
 XX
 SQ Sequence 822 BP; 219 A; 219 C; 217 G; 167 T; 0 other;

Query Match 36.2%; Score 302; DB 23; Length 822;

This Page Blank (uspto)

Db 181 ATCTGAGCCAGGACAGTCTCTAAGTCTCTGACGCGAGAGGAGGCGCAAGCCTTG 240
 QY 241 AAAGACAGTACAAATTTTGAACATCCAGATGCTTTGATAATGATTTGATGACAGAGCT 300
 Db 241 AAAGACAGTACAAATTTTGAACATCCAGATGCTTTGATAATGATTTGATGACAGAGCT 300
 QY 301 CTGAGAAACATCTGAGAGGGGCAAAACGATGAGAGTGGCGACCTATGATTTTGTGACAC 360
 Db 301 CTGAGAAACATCTGAGAGGGGCAAAACGATGAGAGTGGCGACCTATGATTTTGTGACAC 360
 QY 361 TCAAGTTTACAGAGACCAAGTGTCTACCTCTGACGAGTGTCTTGTGTGAGGAGATC 420
 Db 361 TCAAGTTTACAGAGACCAAGTGTCTACCTCTGACGAGTGTCTTGTGTGAGGAGATC 420
 QY 421 TTGGTGTCTTACAGCCAGAGAGATCCGAGACATGTTCCACTGCGCTCTTGTGTGACAC 480
 Db 421 TTGGTGTCTTACAGCCAGAGAGATCCGAGACATGTTCCACTGCGCTCTTGTGTGACAC 480
 QY 481 GACTCCGACGTCAGGCTGTCTCGAAGAGTCTCCGGACGTCGCCAGAGGAGGACCTG 540
 Db 481 GACTCCGACGTCAGGCTGTCTCGAAGAGTCTCCGGACGTCGCCAGAGGAGGACCTG 540
 QY 541 GAGCAGATTCTGACGACATACACCTTCTGTGAAGCCGCTTCAGAGATTTCTGCTG 600
 Db 541 GAGCAGATTCTGACGACATACACCTTCTGTGAAGCCGCTTCAGAGATTTCTGCTG 600
 QY 601 CCGACAAAGAGTATGCGATGTATCATCCACGAGAGTGAACAATATGTGTTCATC 660
 Db 601 CCGACAAAGAGTATGCGATGTATCATCCACGAGAGTGAACAATATGTGTTCATC 660
 QY 661 AACCTGATCTGACGACATCCAGAGATTTCTGAATGTGACATCTGCAAAATGACACGA 720
 Db 661 AACCTGATCTGACGACATCCAGAGATTTCTGAATGTGACATCTGCAAAATGACACGA 720
 QY 721 GAGAGGTCTCAATGGGCGAGCTACAGAGGAGCTTTTCTGAGCCAGAGGAGACCACTG 780
 Db 721 GAGAGGTCTCAATGGGCGAGCTACAGAGGAGCTTTTCTGAGCCAGAGGAGACCACTG 780
 QY 781 ATGCTGACCTCTGCAAAACGCTCACTTTGAGATTCAGACAGACCCCACTGA 834
 Db 781 ATGCTGACCTCTGCAAAACGCTCACTTTGAGATTCAGACAGACCCCACTGA 834

RESULT 2
 US-10-037-270-546
 / Sequence 546, Application US/10037270
 / Publication No. US20030104529A1
 / GENERAL INFORMATION:
 / APPLICANT: Tang, Y. Tom
 / APPLICANT: Liu, Chenghua
 / APPLICANT: Asundi, Vinod
 / APPLICANT: Zhang, Jie
 / APPLICANT: Ren, Feiyan
 / APPLICANT: Chen, Rui-hong
 / APPLICANT: Zhao, Qing A.
 / APPLICANT: Wehrman, Tom
 / APPLICANT: Xue, Aidong J.
 / APPLICANT: Yang, Yonghong
 / APPLICANT: Wang, Jian-Rui
 / APPLICANT: Zhou, Ping
 / APPLICANT: Ma, Yundong
 / APPLICANT: Wang, Dunrui
 / APPLICANT: Wang, Zhiwei
 / APPLICANT: Tillinghast, John
 / APPLICANT: Drmanac, Radoje T.
 / TITLE OF INVENTION: No. US20030104529A1el Nucleic Acids and
 / TITLE OF INVENTION: Polypeptides
 / FILE REFERENCE: 784CIP28
 / CURRENT APPLICATION NUMBER: US/10/037, 270
 / CURRENT FILING DATE: 2002-01-04
 / PRIOR APPLICATION NUMBER: 09/552,317
 / PRIOR FILING DATE: 2000-04-25
 / PRIOR APPLICATION NUMBER: 09/486,725

/ PRIOR FILING DATE: 2000-01-21
 / NUMBER OF SEQ ID NOS: 1104
 / SOFTWARE: pc_files Version 1.0
 / SEQ ID NO 546
 / LENGTH: 1288
 / TYPE: DNA
 / ORGANISM: Homo sapiens
 / NAME/KEY: CDS
 / LOCATION: (95) .. (928)
 / US-10-037-270-546

Query Match 100.0%; Score 834; DB 14; Length 1288;
 Best Local Similarity 100.0%; Pred No. 1,1e-244;
 Matches 834; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCTTCGCGGAGGAGCGAAGACTGCGAGAGCCCGCGCGAGGCGGACCGTCCGAC 60
 Db 95 ATGCTTCGCGGAGGAGCGAAGACTGCGAGAGCCCGCGCGAGGCGGACCGTCCGAC 154
 QY 61 CAGCGCCCTTCTCTGATTAAGGGGTGAGCGCGGCACTGCGAGGAGAAATCGACCTGTGT 120
 Db 155 CAGCGCCCTTCTCTGATTAAGGGGTGAGCGCGGCACTGCGAGGAGAAATCGACCTGTGT 214
 QY 121 GAGAATCATGAGTGTCTGGGACAGAAAGAGTGAACAGCGGACGCGAAGGTGTC 180
 Db 215 GAGAATCATGAGTGTCTGGGACAGAAAGAGTGAACAGCGGACGCGAAGGTGTC 274
 QY 181 ATCTGAGCCAGAGACAGTCTTCAAGAGTCTTACCGGACAGAGCAAGGCGCAAGCCTTG 240
 Db 275 ATCTGAGCCAGAGACAGTCTTCAAGAGTCTTACCGGACAGAGCAAGGCGCAAGCCTTG 334
 QY 241 AAAGACAGTACAAATTTTGAACATCCAGATGCTTTGATAATGATTTGATGACAGAGCT 300
 Db 335 AAAGACAGTACAAATTTTGAACATCCAGATGCTTTGATAATGATTTGATGACAGAGCT 394
 QY 301 CTGAGAAACATCTGAGAGGGCAAAACGATGAGAGTGGCGACCTATGATTTTGTGACAC 360
 Db 395 CTGAGAAACATCTGAGAGGGCAAAACGATGAGAGTGGCGACCTATGATTTTGTGACAC 454
 QY 361 TCAAGTTTACAGAGACCAAGTGTCTACCTCTGACGAGTGTCTTGTGTGAGGAGATC 420
 Db 455 TCAAGTTTACAGAGACCAAGTGTCTACCTCTGACGAGTGTCTTGTGTGAGGAGATC 514
 QY 421 TTGGTGTCTTACAGCCAGAGAGATCCGAGACATGTTCCACTGCGCTCTTGTGTGACAC 480
 Db 515 TTGGTGTCTTACAGCCAGAGAGATCCGAGACATGTTCCACTGCGCTCTTGTGTGACAC 574
 QY 481 GACTCCGACGTCAGGCTGTCTCGAAGAGTCTCCGGACGTCGCCAGAGGAGGACCTG 540
 Db 575 GACTCCGACGTCAGGCTGTCTCGAAGAGTCTCCGGACGTCGCCAGAGGAGGACCTG 634
 QY 541 GAGCAGATTCTGACGACATACACCTTCTGTGAAGCCGCTTCAGAGATTTCTGCTG 600
 Db 635 GAGCAGATTCTGACGACATACACCTTCTGTGAAGCCGCTTCAGAGATTTCTGCTG 694
 QY 601 CCGACAAAGAGTATGCGATGTATCATCCACGAGAGTGAACAATATGTGTTCATC 660
 Db 695 CCGACAAAGAGTATGCGATGTATCATCCACGAGAGTGAACAATATGTGTTCATC 754
 QY 661 AACCTGATCTGACGACATCCAGAGATTTCTGAATGTGACATCTGCAAAATGACACGA 720
 Db 755 AACCTGATCTGACGACATCCAGAGATTTCTGAATGTGACATCTGCAAAATGACACGA 814
 QY 721 GAGAGGTCTCAATGGGCGAGCTACAGAGGAGCTTTTCTGAGCCAGAGGAGACCACTG 780
 Db 815 GAGAGGTCTCAATGGGCGAGCTACAGAGGAGCTTTTCTGAGCCAGAGGAGACCACTG 874
 QY 781 ATGCTGACCTCTGCAAAACGCTCACTTTGAGATTCAGACAGACCCCACTGA 834
 Db 875 ATGCTGACCTCTGCAAAACGCTCACTTTGAGATTCAGACAGACCCCACTGA 928

```
RESULT 3
US-09-896-522-1
; Sequence 1, Application US/09896522
; Patent No. US20020055161A1
; GENERAL INFORMATION:
; APPLICANT: Glucksmann, Maria A.
; TITLE OF INVENTION: 57658, A NOVEL HUMAN URIDINE KINASE AND
; FILE REFERENCE: 381552001700
; CURRENT APPLICATION NUMBER: US/09/896,522
; CURRENT FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: 60/216,503
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1624
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (94)...(927)
; US-09-896-522-1

Query Match          100.0%; Score 834; DB 9; Length 1624;
Best Local Similarity 100.0%; Pred. No. 1,3e-244;
Matches 834; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCTTCGGCGGGAGCGAAGACTGCGAGAGCCCGCGCCGAGGCGGACCGTCCGAC 60
DB 94 ATGGCTTCGGCGGGAGCGAAGACTGCGAGAGCCCGCGCCGAGGCGGACCGTCCGAC 153
QY 61 CACGGCCCTTCTCTGATAGGGGTGAGCGGGCGGCACTCCGACCGGGAAGTGCATCTGT 120
DB 154 CACGGCCCTTCTCTGATAGGGGTGAGCGGGCGGCACTCCGACCGGGAAGTGCATCTGT 213
QY 121 GAGAAGATCATGAGTGTCTGCGACAGAACGAGGTGGAACAGCGCGGCGGAAAGTGTGTC 180
DB 214 GAGAAGATCATGAGTGTCTGCGACAGAACGAGGTGGAACAGCGCGGCGGAAAGTGTGTC 273
QY 181 ATCTTGAGCCAGGACAGGATTTCTACAAAGTCTTGAACGAGGCAAGGCGGCTTG 240
DB 274 ATCTTGAGCCAGGACAGGATTTCTACAAAGTCTTGAACGAGGCAAGGCGGCTTG 333
QY 241 AAAGGACATGATATTTTGAACATCCAGATGCTTTATATATGATTTGATGACAGACT 300
DB 334 AAAGGACATGATATTTTGAACATCCAGATGCTTTATATATGATTTGATGACAGACT 393
QY 301 CTGAAGAACTGTCGAGGCGCAAAACGGTGAAGGTGCGGACTATGATTTTGTGACACAC 360
DB 394 CTGAAGAACTGTCGAGGCGCAAAACGGTGAAGGTGCGGACTATGATTTTGTGACACAC 453
QY 361 TCAAGTTTACAGAGACACAGGTGTCTACCTGCGGACGTGTCTGTTTGAAGGCTATC 420
DB 454 TCAAGTTTACAGAGACACAGGTGTCTACCTGCGGACGTGTCTGTTTGAAGGCTATC 513
QY 421 TTGGTGTCTTACAGCCGAGAGATCCGGAATGTTTCAACTCGGCTCTTCTGTGACACC 480
DB 514 TTGGTGTCTTACAGCCGAGAGATCCGGAATGTTTCAACTCGGCTCTTCTGTGACACC 573
QY 481 GACTCCGACGTCAAGGCTGTCTCGAAGAGTCTCCGGGACGTGCGCGGCGGAGGAGGACCTG 540
DB 574 GACTCCGACGTCAAGGCTGTCTCGAAGAGTCTCCGGGACGTGCGCGGCGGAGGAGGACCTG 633
QY 541 GAGCAGATTTCTGACGAGTACACCACTTGTGAAGCGGCTTTCGAGAGTTCTGCTTG 600
DB 634 GAGCAGATTTCTGACGAGTACACCACTTGTGAAGCGGCTTTCGAGAGTTCTGCTTG 693
QY 601 CCGAAGAAAGATGATGCGGATGATGATATCCCAAGAGAGTGAACAATATGTTGCTATC 660
DB 694 CCGAAGAAAGATGATGCGGATGATGATATCCCAAGAGAGTGAACAATATGTTGCTATC 753
QY 661 AACCTGATCGTGAGACATCCAGSACATCTGAATGTGATCAATCTGCAATGSCACCGA 720
```

```
|||||
DB 754 AACCTGATCGTGAGACATCCAGSACATTTGAATGTGATCTGCAATGTGACACCGA 813
QY 721 GAGGGTCCAAATGCGGAGCTTACAGCGACCTTTTCTGAGCCAGGAGACCACTGGG 780
DB 814 GAGGGTCCAAATGCGGAGCTTACAGCGACCTTTTCTGAGCCAGGAGACCACTGGG 873
QY 781 ATCTGACCTTCTGCAAGGTCACATTTGAGTCCAGAGAGACCCACTGA 834
DB 874 ATCTGACCTTCTGCAAGGTCACATTTGAGTCCAGAGAGACCCACTGA 927

RESULT 4
US-09-833-381-2048
; Sequence 2048, Application US/09833381
; Patent No. US20020132090A1
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. US20020132090A1el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2048
; LENGTH: 1648
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(1648)
; OTHER INFORMATION: n = A,T,C or G
; US-09-833-381-2048

Query Match          98.7%; Score 823; DB 10; Length 1648;
Best Local Similarity 99.9%; Pred. No. 2.9e-241;
Matches 834; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 ATGGCTTCGGCGGGAGCGAAGACTGCGAGAGCCCGCGCCGAGGCGGACCGTCCGAC 60
DB 84 ATGGCTTCGGCGGGAGCGAAGACTGCGAGAGCCCGCGCCGAGGCGGACCGTCCGAC 143
QY 61 CACGGCCCTTCTCTGATAGGGGTGAGCGGGCGGCACTCCAGCGGGAAGTGCATCTGT 120
DB 144 CACGGCCCTTCTCTGATAGGGGTGAGCGGGCGGCACTCCAGCGGGAAGTGCATCTGT 203
QY 121 GAGAAGATCATGAGTGTCTGCGACAGAACGAGGTGGAACAGCGCGGCGGAAAGTGTGTC 180
DB 204 GAGAAGATCATGAGTGTCTGCGACAGAACGAGGTGGAACAGCGCGGCGGAAAGTGTGTC 263
QY 181 ATCTTGAGCCAGGACAGGATTTCTACAAAGTCTTGAACGAGGCAAGGCGGCTTG 240
DB 264 ATCTTGAGCCAGGACAGGATTTCTACAAAGTCTTGAACGAGGCAAGGCGGCTTG 323
QY 241 AAAGGACATGATATTTTGAACATCCAGATGCTTTATATATGATTTGATGACAGACT 300
DB 324 AAAGGACATGATATTTTGAACATCCAGATGCTTTATATATGATTTGATGACAGACT 383
QY 301 CTGAAGAACTGTCGAGGCGCAAAACGGTGAAGGTGCGGACTATGATTTTGTGACACAC 360
DB 384 CTGAAGAACTGTCGAGGCGCAAAACGGTGAAGGTGCGGACTATGATTTTGTGACACAC 443
QY 361 TCAAGTTTACAGAGACACAGGTGTCTACCTGCGGACGTGTCTGTTTGAAGGCTATC 420
DB 444 TCAAGTTTACAGAGACACAGGTGTCTACCTGCGGACGTGTCTGTTTGAAGGCTATC 503
QY 421 TTGGTGTCTTACAGCCGAGAGATCCGGAATGTTTCAACTCGGCTCTTCTGTGACACC 480
DB 504 TTGGTGTCTTACAGCCGAGAGATCCGGAATGTTTCAACTCGGCTCTTCTGTGACACC 563
QY 481 GACTCCGACGTCAAGGCTGTCTCGAAGAGTCTCCGGGACGTGCGCGGCGGAGGAGGACCTG 540
```

```

RESULT 5
US-09-918-995-30379
Sequence 30379, Application US/09918995
Publication No. US20030073623A1
GENERAL INFORMATION:
APPLICANT: Hyaseq, Inc.
TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
FROM VARIOUS CDNA LIBRARIES
FILE REFERENCE: 20411-756
CURRENT APPLICATION NUMBER: US/09/918,995
CURRENT FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: US/09/235,076
PRIOR FILING DATE: 1999-01-20
NUMBER OF SEQ ID NOS: 38054
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 30379.
LENGTH: 472
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(472)
OTHER INFORMATION: n = A,T,C or G
US-09-918-995-30379

```

OY	384	GGTCTACCCGCGGACGTCGTTCTGTGTTGAGGGACATCTTGCTTCACACGCCAGAGAT	443
Db	328	GGTCTACCCGCGGACGTCGTTCTGTGTTGAGGGACATCTTGCTTCACACGCCAGAGAT	387
OY	444	CCGGGACATGTTCCACCTGCGCCTCTTCCTGTGACACCGACTCCGACGTCAGGCTGTCTCG	503
Db	388	CCGGGACATGTTCCACCTGCGCCTCTTCCTGTGACACCGACTCCGACGTCAGGCTGTATCG	447
OY	504	AAGAGTTCCTCCGGACGTGCGCCGA	528
Db	448	AAGAGTTCCTCCGGACGTGCGCCGA	472

QY	DB	Matches	44.8%	Score 374;	DB 13;	Length 1402;
QY	DB	486;	Conservative	0;	Mismatches 165;	Indels 3; Gaps 1
QY	DB	406	GATCGTGGAGCTCCGCGGCGAGAAAGAGTGACTATGCCAGAAAGCAGTGTCTATCT	185		
QY	DB	186	GAGCCAGAGCAGGTCTCTCAAGAGTCTCTGACGCGCAGAGCAGAAAGCCCTTGAAGG	245		
DB	466	GAGCCAGAGTACGCTTCTCAACGCTGTCTTACCTCCGAGCAGAAAGCCCAAGCCCTGAAGG	525			
QY	246	ACAGTACATTTTGTGACCATCCAGATGCGCTTTGATTAATGATTTGATGACAGAGCTCTGAA	305			

```
Db 526 CCAGTTCACCTTTGACCAACCCGGATGCTTTGACAAATGAACTCATTTCTCAAAAACACTCAA 585
Qy 306 GAACATCTGTGAGGGCAAAAACGTTGAGAGTCCGACCTATGATTTGTGACACACTCAAG 365
Db 586 AGAATCACTGAAAGGAAAAACATCCAGATCCCGTGTATGACTTTGTCTCCATTCGCG 645
Qy 366 GTTACCAAGAACCAACGGTGGTCTACCTCGGAGAGTGGTTCGTTTGAAGGGCAATCTTGGT 425
Db 646 GAAGAGAGAGACAGTACTGTCTATCCCGACAGAGTGGTGTCTTTTGAAGGAACTCTGGC 705
Qy 426 GTTTCACGACGAGAGATCCGGAGCATGTTCCACCTGCGCTCTTCGTGACACCGACTC 485
Db 706 CTCTCACTCCAGAGAGTGAAGACCTGTTCCAGATGAAGCTTTTGTGATACAGATGC 765
Qy 486 CGAGCTCAGGCTCTCTGAAAGTTCCTCGGAGAGTGGC---CGAGGAGAGGACCTTGA 542
Db 766 GGAACACCCGGCTCCAGCAGAGATTTAAGGGACATCAGCAGAGAGGACGGAGATCTTGA 825
Qy 543 GCAGATTCGTGACGCACTACCAACCTTCGTGAAGCCGGCTTCGAGGAGTTCTGCTGCC 602
Db 826 GCGATTTTATCTCAGATCACTTACGTTGCTCAGAGCTGCTTTGAGGAAATTCGCTTGC 885
Qy 603 GACAAAGAGATATGCCGATGTATCATCCACGAGAGTGAACAATATGTTGCCATCAA 662
Db 886 AACAAAGAGATATGCTGATGTATCATCCCTAGAGGTGAGATPATCTGGTGGCCATCAA 945
Qy 663 CCTGATCTGTCAGACATCCAGACATTTGAAATGTTGACATCTGCAATATGCA 716
Db 946 CCTCATCTGTGACAGACATCCAGACATCTGAAATGAGGGCCCTCCAAACGGCA 999
```

RESULT 7

```
US-09-925-300-220
; Sequence 220, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben,
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05888
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 220
; LENGTH: 1310
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-300-220
```

Query Match 43.4%; Score 361.6; DB 10; Length 1310;

Best Local Similarity 74.0%; Pred. No. 4.1e-100; Indels 4; Gaps 2;

```
Matches 484; Conservative 1; Mismatches 165; Indels 4; Gaps 2;
Qy 66 GCCCTTCCTGATAGGGGTGAGCGGCGGCACTGCGACGGGGAAGTCGACCGTGTGTGAAA 125
Db 288 GCCCTTCCTATAGGCGTCA-CGGGGGAACAGCTAGGCGCAAGCTTCCGTGTGTCTAA 346
Qy 126 GATCATGAGTGTCTGGAACAACAAGGTGAAACAGCGGCGAGGAGGTGTCTATCT 185
Db 347 GATCTGTCACCTCTGGGGGAGATGAGTGAAGTATCTATCCAGAAAGGTGTCTATCT 406
Qy 186 GAGCAGAGAGAGTTCACAAAGTCTGACGGGAGACAGACAGGCGCAAGGCTTTGAAGG 245
Db 407 GAGCCAGAGATAGCTTCTTACCGTGTCTTACCTCGAGACAGAGGCCAAAGCCCTGAAGG 466
Qy 246 ACAGTACAAATTTTGAACCATCCAGATGCTTTGATTAATGATTTGATGACAGACTCTGAA 305
```

```
Db 467 CCAGTTCACCTTTGACCAACCCGGATGCTTTGACAAATGAAATCTTCTCAAAAACACTCAA 526
Qy 306 GAACATCTGTGAGGGCAAAAACGTTGAGAGTCCGACCTATGATTTGTGACACACTCAAG 365
Db 527 AGAATCACTGAAAGGAAAAACATCCAGATCCCGTGTATGACTTTGTCTCCATTCGCG 586
Qy 366 GTTACCAAGAACCAACGGTGGTCTACCTCGGAGAGTGGTTCGTTTGAAGGGCAATCTTGGT 425
Db 587 GAAGAGAGAGACAGTACTGTCTATCCCGACAGAGTGGTGTCTTTGAAAGGATCTCTGGC 646
Qy 426 GTTTCACGACGAGAGATCCGGAGCATGTTCCACCTGCGCTCTTCGTGACACCGACTC 485
Db 647 CTCTCACTCCAGAGAGTGAAGACCTGTTCCAGATGAAGCTTTTGTGATACAGATGC 706
Qy 486 CGAGCTCAGGCTCTCTGAAAGTTCCTCGGAGAGTGGC---CGAGGAGAGGACCTTGA 542
Db 707 GGAACACCCGGCTCCAGCAGAGATTTAAGGGACATCAGCAGAGAGGACGGAGATCTTGA 766
Qy 543 GCAGATTCGTGACGCACTACCAACCTTCGTGAAGCCGGCTTCGAGGAGTTCTGCTGCC 602
Db 767 GCGATTTTATCTCAGATCACTTACGTTGCTCAGAGCTGCTTTGAGGAAATTCGCTTGC 826
Qy 603 GACAAAGAGATATGCCGATGTATCATCCACGAGAGTGAACAATATGTTGCCATCAA 662
Db 827 AACAAAGAGATATGCTGATGTATCATCCCTAGAGGTGAGATPATCTGGTGGCCATCAA 886
Qy 663 CCTGATCTGTCAGACATCCAGACATTTGAAATGTTGACATCTGCAATATGCA 716
Db 887 CCTCATCTGTGACAGACATCCAGACATCTGAAATGAGGGCCCTCCAAACGGCA 940
```

RESULT 8

```
US-09-918-995-24042
; Sequence 24042, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hysq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 24042
; LENGTH: 472
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(472)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-24042
```

Query Match 26.5%; Score 221; DB 11; Length 472;

Best Local Similarity 73.0%; Pred. No. 3e-57; Indels 0; Gaps 0;

```
Matches 284; Conservative 0; Mismatches 105; Indels 0; Gaps 0;
Qy 66 GCCCTTCCTGATAGGGGTGAGCGGCGGCACTGCGACGGGGAAGTCGACCGTGTGTGAAA 125
Db 84 GCCCTTCCTTATAGGCGTCA-CGGGGGAACAGCTAGGCGCAAGCTTCCGTGTGTCTAA 143
Qy 126 GATCATGAGTGTCTGGAACAACAAGGTGAAACAGCGGCGAGGAGGTGTCTATCT 185
Db 144 GATCTGTCACCTCTGGGGGAGATGAGTGAAGTATCTATCCAGAAAGGTGTCTATCT 203
Qy 186 GAGCAGAGAGAGTTCACAAAGTCTGACGGGAGACAGAGGCCAAGGCTTTGAAGG 245
Db 204 GAGCCAGAGATAGCTTCTTACCGTGTCTTACCTCGAGACAGAGGCCAAAGCCCTGAAGG 263
Qy 246 ACAGTACAAATTTTGAACCATCCAGATGCTTTGATTAATGATTTGATGACAGACTCTGAA 305
```

Position	Sequence	Position	Sequence
Db	264 CCAGTTCAACTTTGACACCCGGATCCTTTGCAATGAATCATTTCTCAAACTCTCA	323	
Oy	306 GAACATCGTGAAGGCAAAACGGTGGAGGCCGACCTATGATTTTGTGCACTCAAG	365	
Db	324 AGAAATCACTGAAGGAAAAACATCCAGATCCCCGTGTGAAGACTTTGTCTCCATTCCG	383	
Oy	366 GTTACAGAGACACGGTGCTACACCTGCGAGCGTGTTCTGTTGAGGSCACTTGGT	425	
Db	384 GAAGAGGGAACAATTACTGTCTATCCCGACGCTGTCTTTGAAGGATCCTGGC	443	
Oy	426 GTTCTACAGCCGAGAGATCCGGGACATGT	454	
Db	444 CTTCTACTCCCGAGAGTACGAACCTGT	472	

```

RESULT 9
US-10-066-543-1843
; Sequence 1843, Application US/10066543
; Publication No. US20030087818A1
GENERAL INFORMATION:
APPLICANT: Jiang, Yugu
APPLICANT: Pyle, Ruth A.
APPLICANT: Xu, Jiangchun
APPLICANT: Indiraas, Carol Yoseph
APPLICANT: Lodes, Michael J.
APPLICANT: Secretist, Heather
APPLICANT: Carter, Darrick
APPLICANT: Fanger, Gary R.
APPLICANT: Smith, Carole L.
APPLICANT: Durham, Margareta
APPLICANT: Stolk, John A.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
FILE REFERENCE: 210121.563
CURRENT APPLICATION NUMBER: US/10/066,543
CURRENT FILING DATE: 2002-01-31
NUMBER OF SEQ ID NOS: 3417
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1843
LENGTH: 447
TYPE: DNA
ORGANISM: Homo sapiens
US-10-066-543-1843

```

Query Match	25.8%;	Score 214.8;	DB 14;	Length 447;
Best Local Similarity	73.0%;	Pred. No. 2.3e-55;		
Matches 276;	Conservative 0;	Mismatches 102;	Indels 0;	Gaps 0;

QY	66	GGCCCTTCCTGAATAGGGGTAGCGGGGCACTGCGACACCGGAAAGTCCACCGTGTGTAGAA	125
Db	70	GGCTTCCTTAATAGCGGTGAGCGGGGAAACGTACCGGCAATCTTCGTGTGTCTAA	129
OY	126	GATCATGAGTGTCTGGGACAGAACAGGTTGGAAACGCGGACGGGAGGTGGTCAATCCT	185
Db	130	GATCTGTCACTCTTGGGGCAAGATAGGTGGACTATCCGCAGAACAGGTGTCTATCCT	189
QY	186	GAGCCAGGACAGGTTCTACMAAGTCTCTGACGGGACAGCAAGAGCCAGGCTTTGAAAG	245
Db	190	GAGCCAGGATAGCTTCTACCGGTCTCTTACCTCGAGCAGMAAGCCAAAGCCTGAAAGG	249
QY	246	ACAGTACAAATTTTGACCAATCCAGATGCCCTTTGTATTAATGATTGGATGCACAGACTCTGAA	305
Db	250	CCAGTTCAACTTTGACCAACCGGATGCTTTTGCATATGAACATCTCTCAAAACAATCAA	309
QY	306	GAACTTCGTGGAGGGCAAAACGGTGGAGGTGCCGACCTATGATTTTGTGCACACTCAAG	365
Db	310	AGAAATCACTGAAGGGAAAAACAGTCCAGATCCCGCGTATGACTTTGTCTCCATTTCCG	369
QY	366	GTTACCAAGACACCGTGTGTCTACCTCCGCGAAGCTGTCTTGTGTGAAGGCATCTTGGT	425
Db	370	GAAGGAGGACAGTTACTGTCTATCCCGACAGCGTGTGCTCTTTGAAGAAATCTGTGC	429

Qy 426 GTTCTACAGCCAGAGAT 443
 ||||| |||||
 Db 430 CTTTACTCCAGAGAGT 447

```

RESULT 10 0995-23923
US-09-918-995-23923
/ Sequence 23923, Application US/09918995
/ Publication No. US20030073623A1
/ GENERAL INFORMATION:
/   APPLICANT: Hyseq, Inc.
/   TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
/   TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
/   FILE REFERENCE: 20411-756
/   CURRENT APPLICATION NUMBER: US/09/918, 995
/   CURRENT FILING DATE: 2001-07-30
/   PRIOR APPLICATION NUMBER: US/09/235, 076
/   PRIOR FILING DATE: 1999-01-20
/   NUMBER OF SEQ. ID NOS: 38054
/   SOFTWARE: FastSeq for Windows Version 3.0
/   SEQ ID NO 23923
/   LENGTH: 455
/   TYPE: DNA
/   ORGANISM: Homo sapiens
/   FEATURE:
/   NAME/KEY: misc_feature
/   LOCATION: (1)..(455)
/   OTHER INFORMATION: n = A,T,C or G
US-09-918-995-23923

```

Query Match	25.2%	Score 210.4;	DB 11;	Length 455;
Best Local Similarity	72.8%;	Pred. No. 5.1e-54;		
Matches 271; Conservative	0;	Mismatches 101;	Indels 0;	Gaps 0;

QY	66	GGCCCTTCTCGAAGAGGGGTGAGCGGGCGGCACTTCCAGCCGGGAAGTGCACCGTGTGTAGAA	125
Db	84	GGCTTCTCTTAAGAGCGTCAAGCGGGGAAACACTAGCCGAAGTCTTCCGTGTGTCTAA	143
QY	126	GATCATGAGTGTCTGGGACAGAACGAGGTGGAAACGCGGACAGCGAAGTGGTCACTCT	185
Db	144	GATCGTCAGCTCTCTGGGGCAGAAATGAGGTGAACTATCGCCAGAAACAGGTGTCTCTT	203
QY	186	GAGCCAGGACAGGTCTTCAACAGGTCTTGAACGACAGACAGAAAGGCCAAGCTTTGAAAG	245
Db	204	GAGCCAGGATGACTTCTTACCGTGTCTTAACCTCGAGCAGAAAGGCCAAGCCCTGAAGG	263
QY	246	ACAGTACAAATTTTGAACCATCCAGATGCCCTTTGATATATGATTTGATGACACAGACTTGA	305
Db	264	CCAGTTCAACTTTGACCAACCGGATGCCCTTTGACATGAAATCATCTCTCAAAACATTCAA	323
QY	306	GAACATCGTGGAGGCAAAACGGTGAAGGTGCGACTTGTATTTTGTGACACACTCAAG	365
Db	324	AGAAATCACTGAAGGAAACAGTCCAGATCCCCGGTGTATGACTTTGTCTCCCATTTCCG	383
QY	366	GTTACACAGAACCAACGAGTGTCTTACCTTCGCGACGATGTCTCTGTTTGAAGGACATTTG	425
Db	384	GAAAGAGGAGACAGTTACTGTCTATCCGACAGATGGTGTCTTTTGAAGGATCTGTGC	443
QY	426	GTTCTACAGCCA	437
Db	444	CTTCACTCCCA	455

```

RESULT 11
US-10-029-386-15893
; Sequence 15893, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR GENOTYPING
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO

```



```
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 15883
; LENGTH: 187
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR9.1
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.5
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.8
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.6
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.7
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.5
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.7
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.3
; OTHER INFORMATION: EST HUMAN HIT: A1992171.1, EVALUE 1.00e-100
; OTHER INFORMATION: SWISSPROT HIT: P52623, EVALUE 4.00e-26
; OTHER INFORMATION: NT HIT: g114785235, EVALUE 1.00e-100
US-10-029-386-15883
```

```
Query Match 21.9%; Score 183; DB 12; Length 187;
Best Local Similarity 100.0%; Pred. No. 8.8e-46;
Matches 183; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 652 GTTGCCATCAACCTGATGTCGACGACATCCAGACATTTCTGAATGTCATCTGCCAA 711
|
|
|
Db 3 GTTGCCATCAACCTGATGTCGACGACATCCAGACATTTCTGAATGTCATCTGCCAA 62
|
|
|
QY 712 TGGCACCAGAGAGGTCCTCATTTGGCCGAGGACGTCACAGCGACCTTTTCTGAGCCAGGGGAC 771
|
|
|
Db 63 TGGCACCAGAGAGGTCCTCATTTGGCCGAGGACGTCACAGCGACCTTTTCTGAGCCAGGGGAC 122
|
|
|
QY 772 CACCTGGAGATGTCGACCTCTGGCAACGGTCACATTTGGAGTCCAGACAGACCCAC 831
|
|
|
Db 123 CACCTGGAGATGTCGACCTCTGGCAACGGTCACATTTGGAGTCCAGACAGACCCAC 182
|
|
|
QY 832 TGA 834
|
|
|
Db 183 TGA 185
|
|
|
RESULT 12
US-10-029-386-2183
; Sequence 2183, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 2183
; LENGTH: 510
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR9.1
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.5
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.8
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.6
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.7
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.5
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.7
```

```
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.3
; OTHER INFORMATION: NT HIT: g11389252, EVALUE 0.00e+00
; OTHER INFORMATION: EST HUMAN HIT: BF664526.1, EVALUE 0.00e+00
; OTHER INFORMATION: SWISSPROT HIT: P52623, EVALUE 2.00e-25
US-10-029-386-2183
```

```
Query Match 21.9%; Score 183; DB 12; Length 510;
Best Local Similarity 100.0%; Pred. No. 1.3e-45;
Matches 183; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 652 GTTGCCATCAACCTGATGTCGACGACATCCAGACATTTCTGAATGTCATCTGCCAA 711
|
|
|
Db 26 GTTGCCATCAACCTGATGTCGACGACATCCAGACATTTCTGAATGTCATCTGCCAA 85
|
|
|
QY 712 TGGCACCAGAGAGGTCCTCATTTGGCCGAGGACGTCACAGCGACCTTTTCTGAGCCAGGGGAC 771
|
|
|
Db 86 TGGCACCAGAGAGGTCCTCATTTGGCCGAGGACGTCACAGCGACCTTTTCTGAGCCAGGGGAC 145
|
|
|
QY 772 CACCTGGAGATGTCGACCTCTGGCAACGGTCACATTTGGAGTCCAGACAGACCCAC 831
|
|
|
Db 146 CACCTGGAGATGTCGACCTCTGGCAACGGTCACATTTGGAGTCCAGACAGACCCAC 205
|
|
|
QY 832 TGA 834
|
|
|
Db 206 TGA 208
|
|
|
```

```
RESULT 13
US-09-764-877-2804
; Sequence 2804, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005
; CURRENT APPLICATION NUMBER: US/09/764,877
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2804
; LENGTH: 9732
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-877-2804
```

```
Query Match 21.9%; Score 183; DB 10; Length 9732;
Best Local Similarity 100.0%; Pred. No. 4e-45;
Matches 183; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 652 GTTGCCATCAACCTGATGTCGACGACATCCAGACATTTCTGAATGTCATCTGCCAA 711
|
|
|
Db 3901 GTTGCCATCAACCTGATGTCGACGACATCCAGACATTTCTGAATGTCATCTGCCAA 3960
|
|
|
QY 712 TGGCACCAGAGAGGTCCTCATTTGGCCGAGGACGTCACAGCGACCTTTTCTGAGCCAGGGGAC 771
|
|
|
Db 3961 TGGCACCAGAGAGGTCCTCATTTGGCCGAGGACGTCACAGCGACCTTTTCTGAGCCAGGGGAC 4020
|
|
|
QY 772 CACCTGGAGATGTCGACCTCTGGCAACGGTCACATTTGGAGTCCAGACAGACCCAC 831
|
|
|
Db 4021 CACCTGGAGATGTCGACCTCTGGCAACGGTCACATTTGGAGTCCAGACAGACCCAC 4080
|
|
|
QY 832 TGA 834
|
|
|
Db 4081 TGA 4083
|
|
|
```

```
RESULT 14
US-09-764-877-2805
; Sequence 2805, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
```

```
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005
; CURRENT APPLICATION NUMBER: US/09/764,877
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2805
; LENGTH: 19125.
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-877-2805

Query Match      21.9%; Score 183; DB 10; Length 19125;
Best Local Similarity 100.0%; Pred. No. 5,2e-45;
Matches 183; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      652 GTTGCATCAACCTGATGTGCGACATCCAGACATTCGAAATGGTACATCTGCAAA 711
DB      5997 GTTGCCATCAACCTGATGTGCGACATCCAGACATTCGAAATGGTACATCTGCAAA 6056
QY      712 TGGGACCGAGAGGGTCCAAATGGGCGAGCTACAAAGCGACCTTTTCTGAGCCAGGGGAC 771
DB      6057 TGGGACCGAGAGGGTCCAAATGGGCGAGCTACAAAGCGACCTTTTCTGAGCCAGGGGAC 6116
QY      772 CACCCCTGGATGTGACCTCTGGCAAAAGGTACATTTGAGTGCAGAGAGACCCAC 831
DB      6117 CACCCCTGGATGTGACCTCTGGCAAAAGGTACATTTGAGTGCAGAGAGACCCAC 6176
QY      832 TGA 834
DB      6177 TGA 6179

RESULT 15
US-10-251-186-14
; Sequence 14, Application US/10251186
; Publication No. US20030180745A1
; GENERAL INFORMATION:
; APPLICANT: Zhang, Y. Tom
; APPLICANT: Zhang, Y. Tom
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20030180745A1 Nucleic Acids and
; FILE REFERENCE: 789CIP2CDV1
; CURRENT APPLICATION NUMBER: US/10/251,186
; CURRENT FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: 09/665,363
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: 09/574,454
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: 09/519,705
; PRIOR FILING DATE: 2000-03-07
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: pt_fl_genes Version 2.0
; SEQ ID NO 14
; LENGTH: 2058
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (173)..(1909)
US-10-251-186-14

Query Match      21.1%; Score 175.6; DB 12; Length 2058;
Best Local Similarity 58.4%; Pred. No. 4.1e-43;
Matches 326; Conservative 0; Mismatches 229; Indels 3; Gaps 1;
```

```
DB      669 ACAAGTGTGACTGAGCAGACAGAGAAAGCCGCGACAAACAACTTCACTTGAC 728
QY      263 ATCCAGATGCTCTTGTGAATGATTTGATGACAGAGACTCTGAAGACATGTGAGAGCA 322
DB      729 ACCGATGTGCTTGTGACTTGACCTCATTTTCCACCTTCAGAAAGCTGAACAGGGGA 788
QY      323 AAACGGTGAAGTGCAGACCTATGATTTGTGACACACTCAAGTTAACAGAGACCGG 382
DB      789 AGAGTGTCAAGGTGCTCCATTTATGACTTACCCACGACAGCCGGAAGAGACTGAAAA 848
QY      383 TGGTCAACCTGCGGACGTGGTCTGTGTTAGGGCATCTTGTGTTCTACAGCCAGAGA 442
DB      849 CACTGTATGTGCAAAAGTCATCATCTTTGAGGGCATATGAGCTTTGCTGACAAAGAC 908
QY      443 TCCGGGACATGTTCCACCTGCGCTCTGTGTGACACCACTCCGACGTGAGCTGTCTC 502
DB      909 TGTGGAGCTCTGGACATGAGATCTTTGTGACACAGACTCCGACATCCGCTGTAC 968
QY      503 GAAAGTTCCTCCGGGACGTGCG---CCGAGGAGGGGACCTGAGACATTCGACGAGT 559
DB      969 GCGGCTGCGCCCGGACATGATGAGCGCGCGGACATCGAGGTGATCAAGCACT 1028
QY      560 ACAACACCTTGTGAAAGCCGCTTCGAGAGTTCGCTGCCGACAAAGAACTATGCCG 619
DB      1029 ACAACAGTTTGTCAAGCCTCTCTCGACCAATGATCCAGCCCAATGCGCTGGCAG 1088
QY      620 ATGTGATCATCCACAGAGAGTGAACATATGTGCTCCATCAACTGATGTGACAGCA 679
DB      1089 ACATCGTGTCCCGACAGAGGGGAGGCAACGCTGGCCATCGACCTGATTTGTGACAGC 1148
QY      680 TCCAGACATTTGTAATG 697
DB      1149 TGCACAGCCAGCTGAGG 1166
```

Search completed: November 25, 2003, 02:19:00
Job time : 307.209 secs


```

Db      181 ATCTGAGCCAGAGAGGTTCTACAGGTCTCTGAAGGAGAGAGAGGAGGAGGCTTG 240
Qy      241 AAAGAGACATCAATTTTGAACCATCCAGATCCCTTTGATTAATGATGATGACACAGACT 300
Db      241 AAAGAGACATCAATTTTGAACCATCCAGATCCCTTTGATTAATGATGATGACACAGACT 300
Qy      301 CTGAAGAAATCTGAGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
Db      301 CTGAAGAAATCTGAGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
Qy      361 TCAAGTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
Db      361 TCAAGTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
Qy      421 TTGGTGTCTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
Db      421 TTGGTGTCTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
Qy      481 GACTCCGACGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
Db      481 GACTCCGACGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
Qy      541 GAGCAGATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
Db      541 GAGCAGATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
Qy      601 CCGACAAAGAGATGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
Db      601 CCGACAAAGAGATGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
Qy      661 AACCTGATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
Db      661 AACCTGATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
Qy      721 GAGAGGTCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
Db      721 GAGAGGTCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
Qy      781 ATGCTGACCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 834
Db      781 ATGCTGACCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 834

```

RESULT 2

```

US-10-037-270-546
; Sequence 546, Application US/10037270
; Publication No. US20030104529a1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yundong
; APPLICANT: Wang, Duntui
; APPLICANT: Wang, Zhiwei
; APPLICANT: Fillinghaas, John
; APPLICANT: Dimaec, Radoje T.
; TITLE OF INVENTION: No. US20030104529a1el Nucleic Acids and
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/10/037,270
; PRIOR FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725

```

```

; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: pc_fl_genes Version 1.0
; SEQ ID NO 546
; LENGTH: 1288
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (95)..(928)
US-10-037-270-546

```

```

Query Match      100.0%; Score 834; DB 14; Length 1288;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 834; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 ATGGCTTCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 60
Db      95 ATGGCTTCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 154
Qy      61 CAGCGGCTCTCTGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
Db      155 CAGCGGCTCTCTGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 214
Qy      121 GAGAAATCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
Db      215 GAGAAATCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 274
Qy      181 ATCTGAGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
Db      275 ATCTGAGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 334
Qy      241 AAAGACATGACAAATTTTGAACCATCCAGATGCTTGTATATATATATATATATATATAT 300
Db      335 AAAGACATGACAAATTTTGAACCATCCAGATGCTTGTATATATATATATATATATATAT 394
Qy      301 CTGAAGAAATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
Db      395 CTGAAGAAATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 454
Qy      361 TCAAGTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
Db      455 TCAAGTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 514
Qy      421 TTGGTGTCTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
Db      515 TTGGTGTCTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 574
Qy      481 GACTCCGACGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
Db      575 GACTCCGACGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 634
Qy      541 GAGCAGATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
Db      635 GAGCAGATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 694
Qy      601 CCGACAAAGAGATGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
Db      695 CCGACAAAGAGATGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 754
Qy      661 AACCTGATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
Db      755 AACCTGATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 814
Qy      721 GAGAGGTCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
Db      815 GAGAGGTCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 874
Qy      781 ATGCTGACCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 834
Db      875 ATGCTGACCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 928

```

this hole is good

```
RESULT 3
US-09-896-522-1
; Sequence 1, Application US/09896522
; Patent No. US2002055161A1
; GENERAL INFORMATION:
; APPLICANT: Glucksmann, Maria A.
; TITLE OF INVENTION: 57658, A NOVEL HUMAN URIDINE KINASE AND
; FILE REFERENCE: 381552001700
; CURRENT APPLICATION NUMBER: US/09/896,522
; CURRENT FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: 60/216,503
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1624
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (94)...(927)
US-09-896-522-1

Query Match          100.0%; Score 834; DB 9; Length 1624;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 834; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCTTCGGCGGAGGCGAAGACTGCGAGAGCCCGCGCGAGGCGGACCGTCCGAC 60
DB 94 ATGGCTTCGGCGGAGGCGAAGACTGCGAGAGCCCGCGCGAGGCGGACCGTCCGAC 153
QY 61 CAGCGGCGCTTCTCTGATAGGGGTGAGCGCGCTCTCCAGCGGGAAGTGCACCTGTGT 120
DB 154 CAGCGGCGCTTCTCTGATAGGGGTGAGCGCGCTCTCCAGCGGGAAGTGCACCTGTGT 213
QY 121 GAGAAATCATGAGTGTGCTGGGACAGAAAGAGTGGACAGCGCGGCGGGAAGTGTGTC 180
DB 214 GAGAAATCATGAGTGTGCTGGGACAGAAAGAGTGGACAGCGCGGCGGGAAGTGTGTC 273
QY 181 ATCTGAGCCAGGACAGGTTCTACAAAGTCTTGAACGACAGGCGGAGGCGGCTTG 240
DB 274 ATCTGAGCCAGGACAGGTTCTACAAAGTCTTGAACGACAGGCGGAGGCGGCTTG 333
QY 241 AAAGGACAGTACATTTTGAACATCCAGATGCTTTATATATGATTTGATGACAGACT 300
DB 334 AAAGGACAGTACATTTTGAACATCCAGATGCTTTATATATGATTTGATGACAGACT 393
QY 301 CTGAAGAATCATGAGTGTGAGGCGAAGGCGTGGAGGTGCGGACTATGATTTTGAACAC 360
DB 394 CTGAAGAATCATGAGTGTGAGGCGAAGGCGTGGAGGTGCGGACTATGATTTTGAACAC 453
QY 361 TCAAGGTTACAGAGACCAACGATGCTTACCTGCGGACGTGTCTTGTGAGGCGATC 420
DB 454 TCAAGGTTACAGAGACCAACGATGCTTACCTGCGGACGTGTCTTGTGAGGCGATC 513
QY 421 TTGGTGTCTTACAGCCAGGAGATCCGGAATGTTCCACTGCGGCTCTTGTGAGACAC 480
DB 514 TTGGTGTCTTACAGCCAGGAGATCCGGAATGTTCCACTGCGGCTCTTGTGAGACAC 573
QY 481 GACTCCGAGTCAAGGCTGTCTGAAAGTCTCCGGGACGTGGCGGAGGGAAGGACCTG 540
DB 574 GACTCCGAGTCAAGGCTGTCTGAAAGTCTCCGGGACGTGGCGGAGGGAAGGACCTG 633
QY 541 GAGCAGATTCTGACGAGTACCAACCTTCTGTAAGCGGCTTCTGAGAGTTCTGCTG 600
DB 634 GAGCAGATTCTGACGAGTACCAACCTTCTGTAAGCGGCTTCTGAGAGTTCTGCTG 693
QY 601 CCGACAAAGAGTATGCGGATGATATCCACAGAGAGTGGACATATGATGTTGCCATC 660
DB 694 CCGACAAAGAGTATGCGGATGATATCCACAGAGAGTGGACATATGATGTTGCCATC 753
QY 661 AACCTGATGATGACATCTCAGGACATCTGAAATGATGATCTTCAAAATGACACGA 720
```

```
RESULT 4
US-09-833-381-2048
; Sequence 2048, Application US/09833381
; Patent No. US20020132090A1
; GENERAL INFORMATION:
; APPLICANT: (Robison, Keith E.)
; TITLE OF INVENTION: NO. US20020132090A1el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2048
; LENGTH: 1648
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(1648)
; OTHER INFORMATION: n = A,T,C or G
US-09-833-381-2048

Query Match          85.6%; Score 714; DB 10; Length 1648;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 834; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 ATGGCTTCGGCGGAGGCGAAGACTGCGAGAGCCCGCGCGAGGCGGACCGTCCGAC 60
DB 84 ATGGCTTCGGCGGAGGCGAAGACTGCGAGAGCCCGCGCGAGGCGGACCGTCCGAC 143
QY 61 CAGCGGCGCTTCTCTGATAGGGGTGAGCGCGGCTCTCCAGCGGGAAGTGCACCTGTGT 120
DB 144 CAGCGGCGCTTCTCTGATAGGGGTGAGCGCGGCTCTCCAGCGGGAAGTGCACCTGTGT 203
QY 121 GAGAAATCATGAGTGTGCTGGGACAGAAAGAGTGGAAACAGCGGCGGGAAGTGTGTC 180
DB 204 GAGAAATCATGAGTGTGCTGGGACAGAAAGAGTGGAAACAGCGGCGGGAAGTGTGTC 263
QY 181 ATCTGAGCCAGGACAGGTTCTTACAAAGTCTTGAACGACAGGCGGAGAGGCGGCTTG 240
DB 264 ATCTGAGCCAGGACAGGTTCTTACAAAGTCTTGAACGACAGGCGGAGAGGCGGCTTG 323
QY 241 AAAGGACAGTACATTTTGAACATCCAGATGCTTTATATATGATTTGATGACAGACT 300
DB 324 AAAGGACAGTACATTTTGAACATCCAGATGCTTTATATATGATTTGATGACAGACT 383
QY 301 CTGAAGAATCATGAGTGTGAGGCGAAGGCGTGGAGGTGCGGACTATGATTTTGAACAC 360
DB 384 CTGAAGAATCATGAGTGTGAGGCGAAGGCGTGGAGGTGCGGACTATGATTTTGAACAC 443
QY 361 TCAAGGTTACAGAGACCAACGATGCTTACCTGCGGACGTGTCTTGTGAGGCGATC 420
DB 444 TCAAGGTTACAGAGACCAACGATGCTTACCTGCGGACGTGTCTTGTGAGGCGATC 503
QY 421 TTGGTGTCTTACAGCCAGGAGATCCGGAATGTTCCACTGCGGCTCTTGTGAGACAC 480
DB 504 TTGGTGTCTTACAGCCAGGAGATCCGGAATGTTCCACTGCGGCTCTTGTGAGACAC 563
QY 481 GACTCCGAGTCAAGGCTGTCTGAAAGTCTCCGGGACGTGGCGGAGGGAAGGACCTG 540
```

Good clone

Db 564 GACTCCGACGTCAGGCTCTTCGAAGAGTCTCCGGGACGTCGCCGAGGAGGAGGAGCCTG 623
 Qy 541 GAGCAGATTCTGACGACATACCACTTCGTGAAGCCGCTTCGAGAGTTCCTGCTG 600
 Db 624 GAGCAGATTCTGACGACATACCACTTCGTGAAGCCGCTTCGAGAGTTCCTGCTG 683
 Qy 601 ~~CCGCAAAAGAGTATGCGGATGATCATCCAGAGAGTGAACAATATGTTGCCAT~~ 659
 Db 684 CCGCAAAAGAGTATGCGGATGATCATCCAGAGAGTGAACAATATGTTGCCAT 743
 Qy 660 CAACCTGATCGTGACGACATCCAGACATTCGTAATGTCATTCGAAATGGACCG 719
 Db 744 CAACCTGATCGTGACGACATCCAGACATTCGTAATGTCATTCGAAATGGACCG 803
 Qy 720 AGAGGGTCCATAGGGCGGAGCTCAAGCGGACCTTTCTGAGCCAGGGACCACTCG 779
 Db 804 AGAGGGTCCATAGGGCGGAGCTCAAGCGGACCTTTCTGAGCCAGGGACCACTCG 863
 Qy 780 GATGCTGACCTCTGCGAAACGCTGACATTTGGAGTCCGACGACACCCCACTGA 834
 Db 864 GATGCTGACCTCTGCGAAACGCTGACATTTGGAGTCCGACGACACCCCACTGA 918

RESULT 5 US-09-918-995-30379

; Sequence 30379, Application US/09918995
 ; Publication No. US20030073623A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Hyseq, Inc.
 ; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
 ; FROM VARIOUS CDNA LIBRARIES
 ; FILE REFERENCE: 20411-756
 ; CURRENT APPLICATION NUMBER: US/09/918, 995
 ; PRIOR FILING DATE: 2001-07-30
 ; PRIOR APPLICATION NUMBER: US/09/235, 076
 ; NUMBER OF SEQ ID NOS: 38054
 ; SOFTWARE: FASTSEQ for Windows Version 3.0
 ; SEQ ID NO 30379
 ; LENGTH: 472
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc.feature
 ; LOCATION: (1)..(472)
 ; OTHER INFORMATION: n = A,T,C or G
 US-09-918-995-30379

date is good

Query Match 48.3%; Score 403; DB 11; Length 472;
 Best Local Similarity 100.0%; Pred. No. 2,4e-207;
 Matches 403; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 97 GCCAGCGGAGTGAACCGTGTGTGAGAGATCATGAGTGTCTGGGACAGAGAGTG 156
 Db 41 GCCAGCGGAGTGAACCGTGTGTGAGAGATCATGAGTGTCTGGGACAGAGAGTG 100
 Qy 157 GAACAGCGGAGCGGAGAGTGTCTGAGCCAGAGCAGGTTCTCAAGTCTTGAAG 216
 Db 101 GAACAGCGGAGCGGAGAGTGTCTGAGCCAGAGCAGGTTCTCAAGTCTTGAAG 160
 Qy 217 GCACAGCAGAGGCGCAAGGCTTGAAGAGACAGTCAATTTTGAACATCCAGAGCCTTT 276
 Db 161 GCACAGCAGAGGCGCAAGGCTTGAAGAGACAGTCAATTTTGAACATCCAGAGCCTTT 220
 Qy 277 GATATGATTTGATGACAGAGCTCTGAAGAAACATTCGTGAGGCGCAAAACGTTGAGAGTG 336
 Db 221 GATATGATTTGATGACAGAGCTCTGAAGAAACATTCGTGAGGCGCAAAACGTTGAGAGTG 280
 Qy 337 CCGACCTATGATTTTGTGACACACTCAAGGTTACAGAGACACGCTGCTTACCTTGC 396
 Db 281 CCGACCTATGATTTTGTGACACACTCAAGGTTACAGAGACACGCTGCTTACCTTGC 340

Qy 397 GAGTGGTCTGTTTGAAGGACATCTTGTTCTACAGCAGGAGATCCGGGACATGTC 456
 Db 341 GAGTGGTCTGTTTGAAGGACATCTTGTTCTACAGCAGGAGATCCGGGACATGTC 400
 Qy 457 CACCTGGCCTCTTCTGTGAGACACGACTCCGACGTGAGGCTGT 499
 Db 401 CACCTGGCCTCTTCTGTGAGACACGACTCCGACGTGAGGCTGT 443

RESULT 6 US-10-029-386-15883

; Sequence 15883, Application US/10029386
 ; Publication No. US20030194704A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Penn, Sharon G.
 ; APPLICANT: Rank, David R.
 ; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR GI
 ; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
 ; FILE REFERENCE: AEOMICA-X-2
 ; CURRENT APPLICATION NUMBER: US/10/029, 386
 ; CURRENT FILING DATE: 2001-12-20
 ; NUMBER OF SEQ ID NOS: 34288
 ; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
 ; SEQ ID NO 15883
 ; LENGTH: 187
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; OTHER INFORMATION: MAP TO CHR9.1
 ; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.5
 ; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.8
 ; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.6
 ; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.7
 ; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3
 ; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.5
 ; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.7
 ; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.3
 ; OTHER INFORMATION: EST HUMAN HIT: A1992171.1, EVALUE 1.00e-100
 ; OTHER INFORMATION: SWISSPROT HIT: P52623, EVALUE 4.00e-26
 ; OTHER INFORMATION: NT HIT: g114783235, EVALUE 1.00e-100
 US-10-029-386-15883

Query Match 21.9%; Score 183; DB 12; Length 187;
 Best Local Similarity 100.0%; Pred. No. 2,4e-88;
 Matches 183; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 652 GTTGCAATCAACTGATCGTGACAGACATCCAGACATTCGTAATGTCATCTGCAAA 711
 Db 3 GTTGCAATCAACTGATCGTGACAGACATCCAGACATTCGTAATGTCATCTGCAAA 62
 Qy 712 TGGCAGCAGAGAGGTTCCAAATGGGCGAGCTACAGCGGACCTTTCTGAGCCAGGGGAC 771
 Db 63 TGGCAGCAGAGAGGTTCCAAATGGGCGAGCTACAGCGGACCTTTCTGAGCCAGGGGAC 122
 Qy 772 CACCTGGGATGCTGACCTCTGGCAACGCTGCAATTTGAGATCCAGACAGACCCAC 831
 Db 123 CACCTGGGATGCTGACCTCTGGCAACGCTGCAATTTGAGATCCAGACAGACCCAC 182
 Qy 832 TGA 834
 Db 183 TGA 185

RESULT 7 US-10-029-386-2183

; Sequence 2183, Application US/10029386
 ; Publication No. US20030194704A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Penn, Sharon G.
 ; APPLICANT: Rank, David R.
 ; APPLICANT: Hanzel, David K.
 ; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR GI

TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
FILE REFERENCE: AEMICA-X-2
CURRENT APPLICATION NUMBER: US/10/029,386
CURRENT FILING DATE: 2001-12-20
NUMBER OF SEQ ID NOS: 34288
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 2183
LENGTH: 510
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO CHR9.1
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.5
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.8
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.6
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.7
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.5
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.7
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.3
OTHER INFORMATION: NT HIT: g113899252, EVALUE 0.00e+00
OTHER INFORMATION: EST HUMAN HIT: BF664526.1, EVALUE 0.00e+00
OTHER INFORMATION: SWISSPROT HIT: P52623, EVALUE 2.00e-25
US-10-029-386-2183

Query Match 21.9%; Score 183; DB 12; Length 510;
Best Local Similarity 100.0%; Pred. No. 2.3e-88;
Matches 183; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 652 GTTGCCATCAACCTGATCGTGACGACATCCAGACATTTGATGTGACATCTGCAAA 711
DB GTTGCCATCAACCTGATCGTGACGACATCCAGACATTTGATGTGACATCTGCAAA 85
QY 712 TGGCACCAGAGAGGCTGCTCAATGGCGGAGCTACAGGCGACCTTTTGTAGCCAGGGAC 771
DB TGGCACCAGAGAGGCTGCTCAATGGCGGAGCTACAGGCGACCTTTTGTAGCCAGGGAC 145
QY 772 CACCTGGGATGCTGACCTCTGGCAACGGTCACTTTGAGTCCAGACAGACCCAC 831
DB CACCTGGGATGCTGACCTCTGGCAACGGTCACTTTGAGTCCAGACAGACCCAC 205
QY 832 TGA 834
DB 206 TGA 208

RESULT 8
US-09-764-877-2804
Sequence 2804, Application US/09764877
Patent No. US20020147140A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PC005
CURRENT APPLICATION NUMBER: US/09/764,877
CURRENT FILING DATE: 2001-01-17
Prior application data removed - refer to PALM or file wrapper
NUMBER OF SEQ ID NOS: 4031
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 2804
LENGTH: 9732
TYPE: DNA
ORGANISM: Homo sapiens
US-09-764-877-2804

Query Match 21.9%; Score 183; DB 10; Length 9732;
Best Local Similarity 100.0%; Pred. No. 2.1e-88;
Matches 183; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 652 GTTGCCATCAACCTGATCGTGACGACATCCAGACATTTGATGTGACATCTGCAAA 711
DB GTTGCCATCAACCTGATCGTGACGACATCCAGACATTTGATGTGACATCTGCAAA 3901

QY 712 TGGCACCAGAGAGGCTCAATGGCGGAGCTACAGCGGACCTTTTCTGAGCCAGGGAC 771
DB TGGCACCAGAGAGGCTCAATGGCGGAGCTACAGCGGACCTTTTCTGAGCCAGGGAC 4020
QY 772 CACCTGGGATGCTGACCTCTGGCAACGGTCACTTTGAGTCCAGACAGACCCAC 831
DB CACCTGGGATGCTGACCTCTGGCAACGGTCACTTTGAGTCCAGACAGACCCAC 4080
QY 832 TGA 834
DB 4081 TGA 4083

RESULT 9
US-09-764-877-2805
Sequence 2805, Application US/09764877
Patent No. US20020147140A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PC005
CURRENT APPLICATION NUMBER: US/09/764,877
CURRENT FILING DATE: 2001-01-17
Prior application data removed - refer to PALM or file wrapper
NUMBER OF SEQ ID NOS: 4031
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 2805
LENGTH: 19125
TYPE: DNA
ORGANISM: Homo sapiens
US-09-764-877-2805

Query Match 21.9%; Score 183; DB 10; Length 19125;
Best Local Similarity 100.0%; Pred. No. 2.1e-88;
Matches 183; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 652 GTTGCCATCAACCTGATCGTGACGACATCCAGACATTTGATGTGACATCTGCAAA 711
DB GTTGCCATCAACCTGATCGTGACGACATCCAGACATTTGATGTGACATCTGCAAA 6056
QY 712 TGGCACCAGAGAGGCTCAATGGCGGAGCTACAGCGGACCTTTTCTGAGCCAGGGAC 771
DB TGGCACCAGAGAGGCTCAATGGCGGAGCTACAGCGGACCTTTTCTGAGCCAGGGAC 6116
QY 772 CACCTGGGATGCTGACCTCTGGCAACGGTCACTTTGAGTCCAGACAGACCCAC 831
DB CACCTGGGATGCTGACCTCTGGCAACGGTCACTTTGAGTCCAGACAGACCCAC 6176
QY 832 TGA 834
DB 6177 TGA 6179

RESULT 10
US-09-908-975-24159
Sequence 24159, Application US/09908975
Publication No. US20030165843A1
GENERAL INFORMATION:
APPLICANT: SHOSHAN, Avi
APPLICANT: WASSERMAN, Alon
APPLICANT: MINTZ, Eli
APPLICANT: MINTZ, Ilat
APPLICANT: FAIGLER, Shmchon
TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICING
FILE REFERENCE: 36688-0005
CURRENT APPLICATION NUMBER: US/09/908,975
CURRENT FILING DATE: 2001-07-20
Prior application data removed - refer to PALM or file wrapper
NUMBER OF SEQ ID NOS: 32337

```
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 24159
; LENGTH: 65
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-908-975-24159
```

```
Query Match          3.1%; Score 26; DB 12; Length 65;
Best Local Similarity 100.0%; Pred. No. 0.0021;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      493 AGGCTGTCTCGAAGAGTTCTCCGGA 518
          |||||
Db       16 AGGCTGTCTCGAAGAGTTCTCCGGA 41
```

```
RESULT 11
US-09-925-300-220
; Sequence 220, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 220
; LENGTH: 1310
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-300-220
```

```
Query Match          2.8%; Score 23; DB 10; Length 1310;
Best Local Similarity 100.0%; Pred. No. 0.079;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      667 ATCGTGACGACATCCAGACAT 689
          |||||
Db       891 ATCGTGACGACATCCAGACAT 913
```

```
RESULT 12
US-10-098-841-53
; Sequence 53, Application US/10098841
; Publication No. US20020197679A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Xu, Chongjun
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yungqing
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhao, Qing A.
; APPLICANT: Ren, Feiyun
; APPLICANT: Chen, Rui-hong
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: Wehtman, Tom
; APPLICANT: Zhang, Jie
; APPLICANT: Qian, Xiaohong B.
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: POLYPEPTIDES
; FILE REFERENCE: 784CIP2
; CURRENT APPLICATION NUMBER: US/10/098,841
```

```
; CURRENT FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: 09/598,042
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 331
; SOFTWARE: pt_fl_genes Version 1.0
; SEQ ID NO 53
; LENGTH: 1402
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (290)..(1075)
US-10-098-841-53
```

```
Query Match          2.8%; Score 23; DB 13; Length 1402;
Best Local Similarity 100.0%; Pred. No. 0.079;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      667 ATCGTGACGACATCCAGACAT 689
          |||||
Db       950 ATCGTGACGACATCCAGACAT 972
```

```
RESULT 13
US-10-066-543-1843
; Sequence 1843, Application US/10066543
; Publication No. US20030087818A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuguo
; APPLICANT: Eyle, Ruth A.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Indrias, Carol Yoseph
; APPLICANT: Lodes, Michael J.
; APPLICANT: Carter, Heather
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Smith, Carole L.
; APPLICANT: Durham, Margarita
; APPLICANT: Stolk, John A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.563
; CURRENT APPLICATION NUMBER: US/10/066,543
; CURRENT FILING DATE: 2002-01-31
; NUMBER OF SEQ ID NOS: 3417
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1843
; LENGTH: 447
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-066-543-1843
```

```
Query Match          2.6%; Score 22; DB 14; Length 447;
Best Local Similarity 100.0%; Pred. No. 0.28;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      173 AGGTGTCATCTGAGCAGGA 194
          |||||
Db       177 AGGTGTCATCTGAGCAGGA 198
```

```
RESULT 14
US-09-918-995-23923
; Sequence 23923, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hysed, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
```



```

; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 23923
; LENGTH: 455
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(455)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-23923

```

```

Query Match          2.6%; Score 22; DB 11; Length 455;
Best Local Similarity 100.0%; Pred. No. 0.28;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      173 AGGTGTCATCTGAGCCAGGA 194
      |||||||||||||||||||
Db      191 AGGTGTCATCTGAGCCAGGA 212

```

```

RESULT 15
US-09-918-995-24042
; Sequence 24042, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hvaeq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 24042
; LENGTH: 472
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(472)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-24042

```

```

Query Match          2.6%; Score 22; DB 11; Length 472;
Best Local Similarity 100.0%; Pred. No. 0.28;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      173 AGGTGTCATCTGAGCCAGGA 194
      |||||||||||||||||||
Db      191 AGGTGTCATCTGAGCCAGGA 212

```

Search completed: November 25, 2003, 07:27:17
 Job time : 675.367 secs

This Page Blank (uspto)

RESULT	1
BM467984	
LOCUS	
DEFINITION	BM467984 1118 bp mRNA linear EST-05-FEB-2002
ACCESSION	AGNCOURT_6437937 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:532885
VERSION	5' , mRNA sequence.
KEYWORDS	BM467984 GI:18517026
SOURCE	EST.
ORGANISM	Homo sapiens (human)
TITLE	Homo sapiens Eukaryota; Metazoa; Chordata; Craniota; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo. 1 (bases 1 to 1118) NIH-MGC http://mgc.nci.nih.gov/. National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished Contact: Robert Strausberg, Ph.D. Email: cgabbs-r@mail.nih.gov Tissue Procurement: ATCC CDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: http://image.lnl.nhl.gov Plate: LLMF2216 row: n column: 22 High quality sequence stop: 689.

FEATURES
source location/Qualifiers
1. 1118
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5532885"
/issue_type="leiomyosarcoma"
/lab_host="DH10B (phage-resistance)"
/clone_id="NIH_MGC_71"
/note="Organ: uterus; Vector: pCMV-Sport6; Site: 1; NotI;
Site: 2; SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2.1 kb."
BASE COUNT 234 a 311 c 361 g 209 t 3 others
ORIGIN
Query Match 96.5%; Score 805.2; DB 12; Length 1118;
Best Local Similarity 98.9%; Pred. No. 1.3e-188;
Matches 821; Conservative 0; Mismatches 8; Indels 1; Gaps 1;
QY 1 ATGGCTTCGGGGGAGGCGAAGACTGCGAGAGCCCGCGGAGCGGACCGTCCGAC 60
DB 47 ATGGCTTCGGGGGAGGCGAAGACTGCGAGAGCCCGCGGAGCGGACCGTCCGAC 106
QY 61 CAGCGCCCTTCCTGATAGGGGTGAGCGCGCACTGCCAGCGGAAAGTCAGCCGTGT 120
DB 107 CAGCGCCCTTCCTGATAGGGGTGAGCGCGCACTGCCAGCGGAAAGTCAGCCGTGT 166
QY 121 GAGAAATCATGAGTTCCTGGGACAGAACGAGGTGAAACGCGCAGCGGAGGTGTC 180
DB 167 GAGAAATCATGAGTTCCTGGGACAGAACGAGGTGAAACGCGCAGCGGAGGTGTC 226
QY 181 ATCTGAGCCAGGACAGGTTCTACAGGTCCTGACGCGAGGAGGAGGCGGACCTTG 240
DB 227 ATCTGAGCCAGGACAGGTTCTACAGGTCCTGACGCGAGGAGGAGGCGGACCTTG 286
QY 241 AAAGACAGTACATTTTGAACATCCAGATGCTTTGATATGATTTGATGACAGGACT 300
DB 287 AAAGACAGTACATTTTGAACATCCAGATGCTTTGATATGATTTGATGACAGGACT 346
QY 301 CTGAAGAACTGTCGAGGCGCAAAACGCTGAGGTCGCCAGCTATGATTTTGTGACAC 360
DB 347 CTGAAGAACTGTCGAGGCGCAAAACGCTGAGGTCGCCAGCTATGATTTTGTGACAC 406
QY 361 TCAAGTTACAGAGACCAAGGTGTCTACCCCTGCGGACGCTGTTGTTGAGGGATC 420
DB 407 TCAAGTTACAGAGACCAAGGTGTCTACCCCTGCGGACGCTGTTGTTGAGGGATC 466
QY 421 TTGGTGTCTACAGCCAGAGATCCGGGACATGTTCCACCTGCGCTCTTCTGTCGACAC 480
DB 467 TTGGTGTCTACAGCCAGAGATCCGGGACATGTTCCACCTGCGCTCTTCTGTCGACAC 526
QY 481 GACTCCGACGTCAAGCTGTCTCGAAGATTCTCCGGGACGTGCGCCGAGGAGGACCTG 540
DB 527 GACTCCGACGTCAAGCTGTCTCGAAGATTCTCCGGGACGTGCGCCGAGGAGGACCTG 586
QY 541 GAGGAGATTGAGGAGTACAGCACTTGTGTAAGCGCGCTTCGAGGAGTTCTGCTG 600
DB 587 GAGGAGATTGAGGAGTACAGCACTTGTGTAAGCGCGCTTCGAGGAGTTCTGCTG 646
QY 601 CCGAACAAGAGTATGCGATGATGATCCACAGAGAGTGAACAATATGTTGCCATC 660
DB 647 CCGAACAAGAGTATGCGATGATGATCCACAGAGAGTGAACAATATGTTGCCATC 706
QY 661 AACTGATTCGTGACGACATCCAGGACATTCGATGATGATGATGATGATGATGATG 720
DB 707 AACTGATTCGTGACGACATCCAGGACATTCGATGATGATGATGATGATGATGATG 766
QY 721 GAGAGGTTCATGAGGCGGAGCTAAGCGGACCTTTCTGAGCGAGGAGGACACCTGG- 779
DB 767 GAGAGGTTCATGAGGCGGAGCTAAGCGGACCTTTCTGAGCGAGGAGGACACCTGG- 826
QY 780 GATGCTACCTTGGCAAAACGCTGACATTTGAGTTCAGGACGAGACCC 829

DB 827 GATGCTACCTTGGCAAGGCTCCACATTTGAGTTCAGGACGAGACCC 876
RESULT 2
BX400889 998 bp mRNA linear EST 13-MAY-2003
LOCUS BX400889 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
DEFINITION CDNA clone CS0DK003YL24 5-PRIME, mRNA sequence.
ACCESSION BX400889
VERSION BX400889.1 GI:30622359
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 998)
AUTHORS Li, W.-B., Gruber, C., Jessee, J. and Polyes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 4968.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DK003YD12Q1&cluster=4968.r Contact :
Feng Liang Email : fliang@life.com/ invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DK003YD12Q1.
FEATURES
source location/Qualifiers
1. 998
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DK003YL24"
/cell_type="HELA CELLS COT 25-NORMALIZED"
/cell_line="HELA"
/clone_id="HELA"
/note="1st strand cDNA was primed with a NotI-Oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT 223 a 263 c 314 g 186 t 12 others
ORIGIN
Query Match 93.7%; Score 781.6; DB 13; Length 998;
Best Local Similarity 98.9%; Pred. No. 8.3e-183;
Matches 825; Conservative 2; Mismatches 3; Indels 4; Gaps 4;
QY 1 ATGGCTTCGGGGGAGGCGAAGACTGCGAGAGCCCGCGGAGCGGACCGTCCGAC 60
DB 67 ATGGCTTCGGGGGAGGCGAAGACTGCGAGAGCCCGCGGAGCGGACCGTCCGAC 125
QY 61 CAGCGCCCTTCCTGATAGGGGTGAGCGCGCACTGCCAGCGGAGTCAAGCTGTGT 120
DB 126 CAGCGCCCTTCCTGATAGGGGTGAGCGCGCACTGCCAGCGGAGTCAAGCTGTGT 185
QY 121 GAGAAATCATGAGTTCCTGGGACAGAACGAGGTGAAACAGCGGACGCGGAGGTC 180
DB 186 GAGAAATCATGAGTTCCTGGGACAGAACGAGGTGAAACAGCGGACGCGGAGGTC 245
QY 181 ATCTGAGCCAGGACAGGTTCTACAGAGTCTCGACGCGAGAGCAAGAGCCCTTG 240
DB 246 ATCTGAGCCAGGACAGGTTCTACAGAGTCTCGACGCGAGAGCAAGAGCCCTTG 305
QY 241 AAAGACAGTACATTTTGAACATCCAGATGCTTTGATATGATTTGATGACAGGACT 300
DB 306 AAAGACAGTACATTTTGAACATCCAGATGCTTTGATATGATTTGATGACAGGACT 365
QY 301 CTGAAGAACTGTCGAGGCGCAAAACGCTGAGGTCGCCAGCTATGATTTTGTGACAC 360

Db 366 CTGAAGACATCGTGGAGGGCAAAACGTTGAGTGCCGACCTATGATTTTGTGACAC 425
 Qy 361 TCAAGTTACAGAGACAGGCTGTACCTGCGGACCTGCTTCTTTGAGGGATC 420
 Db 426 TCAAGTTACAGAGACAGGCTGTACCTGCGGACCTGCTTCTTTGAGGGATC 485
 Qy 421 TTGGTGTCTACAGCCAGAGATCCGGACATGTTCCAGCTGCGCTCTTGTGAGAC 480
 Db 486 TTGGTGTCTACAGCCAGAGATCCGGACATGTTCCAGCTGCGCTCTTGTGAGAC 545
 Qy 481 GACTCCGACGTGAGCTGTCTCGAAGATTCTCGGACGTGCGCGGAGAGGACCTG 540
 Db 546 GACTCCGACGTGAGCTGTCTCGAAGATTCTCGGACGTGCGCGGAGAGGACCTG 605
 Qy 541 GAGCAATTTTGAAGCGATCACCACTTGTGAGGCGGCTTTCCAGAGATTCTGCTG 600
 Db 606 GAGCAATTTTGAAGCGATCACCACTTGTGAGGCGGCTTTCCAGAGATTCTGCTG 664
 Qy 601 CCGACAAAGAGATGCGGATGATCATCCAGAGAGTGGACAAATATGTTGGCATC 660
 Db 665 CCGACAAAGAGATGCGGATGATCATCCAGAGAGTGGACAAATATGTTGGCATC 724
 Qy 661 AACCTGATGTCAGCAATCCAGACATTTCTGAATGTGACATCTGCAATGACACGA 720
 Db 725 AACCTGATGTCAGCAATCCAGACATTTCTGAATGTGACATCTGCAATGACACGA 784
 Qy 721 GAGGGTTCCTAATGGCGGAGCTACAGCGACCTTTTCTGAGCGAGGGAGCCCTGG 780
 Db 785 GAGGGTTCCTAATGGCGGAGCTACAGCGACCTTTTCTGAGCGAGGGAGCCCTGG 842
 Qy 781 ATCTGACCTCTGAGCAACGCTGACATTTGAGTCCAGAGACCCCACTGA 834
 Db 843 ATCTGACCTCTGAGCAACGCTGACATTTGAGTCCAGAGACCCCACTGA 896

RESULT 3
 LOCUS BX394295 1201 bp mRNA linear EST 13-MAY-2003
 DEFINITION BX394295 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
 CDNA clone CS0DC013YE13 5-PRIME, mRNA sequence.
 ACCESSION BX394295
 VERSION BX394295.1 GI:30624219
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 Li, W.B., Gruber, C., Jesssee, J. and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: sequef@genoscope.cns.fr, Web: www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 4968.r For more information about this cluster, see
 http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CS0DC013AC070P1&cluster=4968.r. Contact :
 Feng Liang Email: fliang@life.techn.com URL :
 http://fulllength.invitrogen.com/Invitrogen Corporation 1600
 Faraday Avenue Genoscope Sequence ID : CS0DC013AC070P1.
 Location/Qualifiers
 1. 1201
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DC013YE13"
 /tissue_type="NEUROBLASTOMA COT 25-NORMALIZED"
 /clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was

FEATURES

source
 1. 1201
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DC013YE13"
 /tissue_type="NEUROBLASTOMA COT 25-NORMALIZED"
 /clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was

digested with Not I and cloned into the Not I and EcoR V
 sites of the pCMVSPORT 6 vector. Library was normalized."
 BASE COUNT 288 a 284 c 363 g 226 t 40 others
 ORIGIN
 Query Match 93.4%; Score 779; DB 13; Length 1201;
 Best Local Similarity 99.4%; Pred. No. 4e-182;
 Matches 823; Conservative 0; Mismatches 1; Indels 4; Gaps 4;

Qy 1 ATGGCTTGGCGGAGGCGAAGACTGCGAGAGCCCGCGAGGCCGACCGTCCGAC 60
 Db 88 ATGGCTTGGCGGAGGCGAAGACTGCGAGAGCCCGCGAGGCCGACCGTCCGAC 146
 Qy 61 CAGCGGCCCTTCTCTGTAAGGGGTGAGCGGCGGACCTGCGAGGGAGAGTGCACCGTGT 120
 Db 147 CAGCGGCCCTTCTCTGTAAGGGGTGAGCGGCGGACCTGCGAGGGAGAGTGCACCGTGT 206
 Qy 121 GAGAAATCATGAGATTGCTGGGACAGAACGAGTGGAAACAGCGGACGGAAGTGTG 180
 Db 207 GAGAAATCATGAGATTGCTGGGACAGAACGAGTGGAAACAGCGGACGGAAGTGTG 266
 Qy 181 ATCTGAGCCAGGACAGGTTCTACAAAGTCTGACGCGCAGACAGAAAGGCAAGGCTTTG 240
 Db 267 ATCTGAGCCAGGACAGGTTCTACAAAGTCTGACGCGCAGACAGAAAGGCAAGGCTTTG 326
 Qy 241 AAAGCAGTAACTTTTGAACATCCAGATGCGCTTTGATTAATGATTTGATGACAGACT 300
 Db 327 AAAGCAGTAACTTTTGAACATCCAGATGCGCTTTGATTAATGATTTGATGACAGACT 386
 Qy 301 CTGAAGACATGCTGGAGGGCAAAACGGTGAAGTGCAGACCTATGATTTTGTGACAC 360
 Db 387 CTGAAGACATGCTGGAGGGCAAAACGGTGAAGTGCAGACCTATGATTTTGTGACAC 446
 Qy 361 TCAAGTTACAGAGACACGCTGTCTACAAAGTCTGACGCGCAGACAGAAAGGCAAGGCTTTG 420
 Db 447 TCAAGTTACAGAGACACGCTGTCTACAAAGTCTGACGCGCAGACAGAAAGGCAAGGCTTTG 506
 Qy 421 TTGGTGTCTTACAGCAGAGAGATCCGGGACATGTTCCACCTGCGCTCTTGTGACAC 480
 Db 507 TTGGTGTCTTACAGCAGAGAGATCCGGGACATGTTCCACCTGCGCTCTTGTGACAC 566
 Qy 481 GACTCCGACGTGAGCTGTCTCGAAGATTCTCGGAGAGCTGCGCGGAGAGGAGACTG 540
 Db 567 GACTCCGACGTGAGCTGTCTCGAAGATTCTCGGAGAGCTGCGCGGAGAGGAGACTG 626
 Qy 541 GAGCAGATTTCTGACGCAATCAACACTTGTGAAAGCCGCGCTTGTGAGAGATTCTGCTG 600
 Db 627 GAGCAGATTTCTGACGCAATCAACACTTGTGAAAGCCGCGCTTGTGAGAGATTCTGCTG 686
 Qy 601 CCGACAAAGAGATGCGGATGATCATCCAGAGAGTGGACAAATATGTTGGCATC 660
 Db 687 CCGACAAAGAGATGCGGATGATCATCCAGAGAGTGGACAAATATGTTGGCATC 746
 Qy 661 AACCTGATGTCAGCAATCCAGACATTTCTGAATGTGACATCTGCAATGACACGA 720
 Db 747 AACCTGATGTCAGCAATCCAGACATTTCTGAATGTGACATCTGCAATGACACGA 806
 Qy 721 GAGGGTTCCTAATGGCGGAGCTACAGCGACCTTTTCTGAGCGAGGGAGCAACCTGGG 780
 Db 807 GAGGGTTCCTAATGGCGGAGCTACAGCGACCTTTTCTGAGCGAGGGAGCAACCTGGG 864
 Qy 781 ATGCTGACCTCTGCGAAAGGCTACATTTGAGTCCAGAGACAGACCC 828
 Db 865 ATGCTGACCTCTGCGAAAGGCTACATTTGAGTCCAGAGACAGACCC 911

RESULT 4
 LOCUS BX343101 1114 bp mRNA linear EST 02-MAY-2003
 DEFINITION BX343101 Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED
 Homo sapiens cDNA clone CS0DL009YD15 5-PRIME, mRNA sequence.
 ACCESSION BX343101
 VERSION BX343101.1 GI:30334170

KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS 1 (bases 1 to 1114)
TITLE L.J. B., Gruber, C., Jesse, J. and Polayes, D.
JOURNAL Pull-length cDNA libraries and normalization
COMMENT Unpublished
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 4968.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DL009CB080P1&cluster=4968.r. Contact :
Peng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paradey Avenue Genoscope sequence ID : CS0DL009CB080P1.
FEATURES
source
1. 1114
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DL0097D15"
/cell_type="B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"
/cell_line="RAMOS CELL LINE"
/clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE) COT
25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo (4T)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT 247 a 291 c 360 g 202 t 14 others
ORIGIN
Query Match 92.9%; Score 774.4; DB 13; Length 1114;
Best Local Similarity 98.9%; Pred. No. 5.3e-181;
Matches 819; Conservative 1; Mismatches 4; Indels 4; Gaps 4;
QY 1 ATGGCTTCGGCGGAGGCGAAGACTGCGAGAGCCCGCGCGGACCGTCCGAC 60
DB 159 ATGGCTTCGGCGGAGGCGAAGACTGCGAGAGCCCGCGCGGACCGTCCGAC 217
QY 61 CAGCGGCCCTTCTGATAGGGGTGAGCGCGGCACTGCGAGCGGAAAGTCCGCTGT 120
DB 218 CAGCGGCCCTTCTGATAGGGGTGAGCGCGGCACTGCGAGCGGAAAGTCCGCTGT 277
QY 121 GAGAGATCATGAGATTCTGGGACAGAACAGAGTGGAAACGCGGACCGGAAGTGTGTC 180
DB 278 GAGAGATCATGAGATTCTGGGACAGAACAGAGTGGAAACGCGGACCGGAAGTGTGTC 337
QY 121 GAGAGATCATGAGATTCTGGGACAGAACAGAGTGGAAACGCGGACCGGAAGTGTGTC 180
DB 278 GAGAGATCATGAGATTCTGGGACAGAACAGAGTGGAAACGCGGACCGGAAGTGTGTC 337
QY 181 ATCTGAGCCGAGACAGGTTCTACAGGTTCTGACGCGAGACAGAGGCCAAGCCCTTG 240
DB 338 ATCTGAGCCGAGACAGGTTCTACAGGTTCTGACGCGAGACAGAGGCCAAGCCCTTG 397
QY 241 AAGAGACAGTACATTTTGAACCATCCAGATCCTTTGATTAATGATTGATGACAGACT 300
DB 398 AAGAGACAGTACATTTTGAACCATCCAGATCCTTTGATTAATGATTGATGACAGACT 457
QY 301 CTGAGAAACATCTGAGAGGGGAAACAGGTGAGAGTGCACCTATGATTTTGTACACAC 360
DB 458 CTGAGAAACATCTGAGAGGGGAAACAGGTGAGAGTGCACCTATGATTTTGTACACAC 517
QY 361 TCAAGTTTACAGAGACACGCGTGTCTACCTCGGACGCGTGTCTGTTGAGGGATC 420
DB 518 TCAAGTTTACAGAGACACGCGTGTCTACCTCGGACGCGTGTCTGTTGAGGGATC 577
QY 421 TTGGTGTCTTACAGCCAGAGATCCGGACATGTTTCCACTCGCCCTTCTGTTGAGACAC 480
DB 578 TTGGTGTCTTACAGCCAGAGATCCGGACATGTTTCCACTCGCCCTTCTGTTGAGACAC 637

QY 481 GACTCCGAGCTCAGGCTGTCTCGAAGATTCTCCGGGACGTGCGCGGAGGAGGACCTG 540
DB 638 GACTCCGAGCTCAGGCTGTCTCGAAGATTCTCCGGGACGTGCGCGGAGGAGGACCTG 697
QY 541 GAGCAGATTCTGACGAGTACACCACTTCTGTAAGCCGCGCTTCCAGAGATTCTGCTG 600
DB 698 GAGCAGATTCTGACGAGTACACCACTTCTGTAAGCCGCGCTTCCAGAGATTCTGCTG 757
QY 601 CCGACAAAGAGTATCCGATGATGATCATCCACGAGAGTGGACATATGTTGCCATC 660
DB 758 CCGACAAAGAGTATCCGATGATGATCATCCACGAGAGTGGACATATGTTGCCATC 817
QY 661 AACCTGATGTCGAGACATCCAGACATTTCTGAATGTGACATCTGCAATGGACCGA 720
DB 818 AACCTGATGTCGAGACATCCAGACATTTCTGAATGTGACATCTGCAATGGACCGA 877
QY 721 GAGAGTTCATGAGGCGGAGCTCAAGCGGACCTTTCTGAGCCAGGAGCCACCTGGG 780
DB 878 GAGAGTTCATGAGGCGGAGCTCAAGCGGACCTTTCTGAGCCAGGAGCCACCTGGG 935
QY 781 ATCTGACCTCTGCGCAACGCTGCACTTTGAGTCCAGACGACACC 828
DB 936 ATCTGACCTCTGCGCAACGCTGCACTTTGAGTCCAGACGACACC 982
RESULT 5
B0719741 956 bp mRNA linear EST 16-JUL-2002
AGENCOURT 8229663 lupski dorsal root ganglion Homo sapiens cDNA
clone IMAGE:6184625 5', mRNA sequence.
B0719741
B0719741.1 GI:21858638
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 956)
NIH-MGC http://mgc.ncl.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgabs@mail.nih.gov
Tissue Procurement: Dr. James R. Lupski
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLMK13574 row: b column: 18
High quality sequence stop: 552.
Location/Qualifiers
1. 956
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6184625"
/sex="male"
/tissue_type="dorsal root ganglia"
/dev stage="adult, 36 yr"
/lab_host="DH10B"
/clone_lib="Lupski dorsal root ganglion"
/note="Vector: pCMV-SPORT6 (Life Technologies); Site 1:
NotI; Site 2: SalI; cDNA made by oligo-dT priming.
Directionally cloned using the following adaptors:
5'-TGACCCAGCGTCCG-3' and
5'-GACTGTTCTTAATCCGAGCGGCGCCCT(15)-3'. Size selected >
1 kb for average insert length 1.7 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
College of Medicine) and is available through Life

Technologies."

BASE COUNT 202 a 272 c 296 g 186 t

ORIGIN

Query Match 90.0%; Score 750.4; DB 13; Length 956;
 Best Local Similarity 95.9%; Pred. No. 4.3e-175;
 Matches 792; Conservative 0; Mismatches 31; Indels 3; Gaps 2;

QY 1 ATGGCTTGGCGGGAGGCGAAGACTGCGAGAGCCCCCGCGGAGGCGGACCGTCCGAC 60
 DB 7 ATGGCTTGGCGGGAGGCGAAGACTGCGAGAGCCCCCGCGGAGGCGGACCGTCCGAC 66
 QY 61 CAGCGGCGCTTCTGATAGGGGTGAGCGGCGGCACTGCGAGCGGGAAGTGCACCGTGT 120
 DB 67 CAGCGGCGCTTCTGATAGGGGTGAGCGGCGGCACTGCGAGCGGGAAGTGCACCGTGT 126
 QY 121 GAGAAATCATGAGATTCTGCGGACAGAAAGAGGTGGAAGCGGCGAGCGGAAGTGTGTC 180
 DB 127 GAGAAATCATGAGATTCTGCGGACAGAAAGAGGTGGAAGCGGCGAGCGGAAGTGTGTC 186
 QY 181 ATCTTGAGCCAGAGCAGGTTCTACAGAGTCTTGAACGCGACAGAAAGGCGGACCTTG 240
 DB 187 ATCTTGAGCCAGAGCAGGTTCTACAGAGTCTTGAACGCGACAGAAAGGCGGACCTTG 246
 QY 241 AAAGSACATGACATTTTGAACATCCAGATGCGCTTGTATGATTTGATGACAGAGCT 300
 DB 247 AAAGSACATGACATTTTGAACATCCAGATGCGCTTGTATGATTTGATGACAGAGCT 306
 QY 301 CTGAAGAATCATGCTGAGGCGAAACCGGTGAGGTGCGACCTATGATTTTGTGACACAC 360
 DB 307 CTGAAGAATCATGCTGAGGCGAAACCGGTGAGGTGCGACCTATGATTTTGTGACACAC 366
 QY 361 TCAAGGTACCAAGACCAAGGTGTCTACACCTGCGGACGTGGTCTGTTTGAAGGATC 420
 DB 367 TCAAGGTACCAAGACCAAGGTGTCTACACCTGCGGACGTGGTCTGTTTGAAGGATC 426
 QY 421 TTGGTGTCTACAGCCAGAGATCCGGGACATGTTCCAGCTGGGCGCTTCTGTTGACAC 480
 DB 427 TTGGTGTCTACAGCCAGAGATCCGGGACATGTTCCAGCTGGGCGCTTCTGTTGACAC 486
 QY 481 GACTCCGAGCTCAGGCTGTCTCGAAGATTCTCCGGACGTGCGCCGAGGGAAGGACTG 540
 DB 487 GACTCCGAGCTCAGGCTGTCTCGAAGATTCTCCGGACGTGCGCCGAGGGAAGGACTG 546
 QY 541 GAGCAGATTCTGACGCGATCACCACTTGTGTAAGCGGCGCTTCAAGGATTTGCTGCTG 600
 DB 547 GAGCAGATTCTGACGCGATCACCACTTGTGTAAGCGGCGCTTCAAGGATTTGCTGCTG 606
 QY 601 CGGACAAAGAGATGTCGATGATGATCCCAAGAGAGTGGACATATAGG--TTGCGCAT 659
 DB 607 CGGACAAAGAGATGTCGATGATGATCCCAAGAGAGTGGACATATATAGTGTGCGCAT 666
 QY 660 CAACCTGATGTCAGACATCCAGGACATTTGTAATGTGATCATCTGCAATATGCGACCG 719
 DB 667 CAACCTGATGTCAGACATCCAGGACATTTTAAATGGGGAACCTGCAAAATGCGACCC 726
 QY 720 AGGAGGTCCTAATGGGCGGAGCTACAGGCGGACCTTTTCTGAGCC--AGGAGACACCT 777
 DB 727 AGGAGGTCCTAATGGGCGGAGCTTCAAGGCGGACCTTTTCTTAAACCCAGGGAACCCCT 786
 QY 778 GGGATGTCGACTCTGCGCAAAAGGTCACTTTGGAGTCCAGCAGCA 823
 DB 787 GGGATGTCGACTCTTGGCTTAAGGTTCAAATTTGGAGGCGCAGCA 832

RESULT 6
 BM459108 1051 bp mRNA linear EST 05-FEB-2002
 LOCUS AGENCOURT_6414543 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5557433
 DEFINITION 5', mRNA Sequence.
 ACCESSION BM459108
 VERSION BM459108.1 GI:18508148
 KEYWORDS EST.

SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 1051)
 AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-rc@mail.nih.gov
 Tissue Procurement: ATCC/DCTD/DP
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/BLN at: <http://image.llnl.gov>
 Plate: LLAM12278 row: m column: 18
 High quality sequence stop: 542.

FEATURES
 source
 1..1051
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone_lib="NIH_MGC_72"
 /note="Organ: skin; Vector: pCMV-SPORT6; Site: 1: NotI; Site: 2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 2 kb. Library constructed by Life Technologies."

BASE COUNT 237 a 264 c 312 g 221 t 17 others

ORIGIN

Query Match 90.0%; Score 750.2; DB 12; Length 1051;
 Best Local Similarity 97.5%; Pred. No. 5e-175;
 Matches 783; Conservative 0; Mismatches 18; Indels 2; Gaps 2;

QY 1 ATGGCTTGGCGGGAGGCGAAGACTGCGAGAGCCCCCGGAGGCGGACCGTCCGAC 60
 DB 6 ATGGCTTGGCGGGAGGCGAAGACTGCGAGAGCCCCCGGAGGCGGACCGTCCGAC 65
 QY 61 CAGCGGCGCTTCTGATAGGGGTGAGCGGCGGCACTGCGAGCGGGAAGTGCACCGTGT 120
 DB 66 CAGCGGCGCTTCTGATAGGGGTGAGCGGCGGCACTGCGAGCGGGAAGTGCACCGTGT 125
 QY 121 GAGAAATCATGAGATTCTGCGGACAGAAAGAGGTGGAAGCGGCGAGCGGAAGTGTGTC 180
 DB 126 GAGAAATCATGAGATTCTGCGGACAGAAAGAGGTGGAAGCGGCGAGCGGAAGTGTGTC 185
 QY 181 ATCTTGAGCCAGAGCAGGTTCTACAGAGTCTTGAACGCGACAGAAAGGCGGACCTTG 240
 DB 186 ATCTTGAGCCAGAGCAGGTTCTACAGAGTCTTGAACGCGACAGAAAGGCGGACCTTG 245
 QY 241 AAAGSACATGACATTTTGAACATCCAGATGCGCTTGTATGATTTGATGACAGAGCT 300
 DB 246 AAAGSACATGACATTTTGAACATCCAGATGCGCTTGTATGATTTGATGACAGAGCT 305
 QY 301 CTGAAGAATCATGCTGAGGCGAAACCGGTGAGGTGCGACCTATGATTTTGTGACACAC 360
 DB 306 CTGAAGAATCATGCTGAGGCGAAACCGGTGAGGTGCGACCTATGATTTTGTGACACAC 365
 QY 361 TCAAGGTACCAAGACCAAGGTGTCTACACCTGCGGACGTGGTCTGTTTGAAGGATC 420
 DB 366 TCAAGGTACCAAGACCAAGGTGTCTACACCTGCGGACGTGGTCTGTTTGAAGGATC 425
 QY 421 TTGGTGTCTACAGCCAGAGATCCGGGACATGTTCCAGCTGCGCTTCTGTTGACAC 480
 DB 426 TTGGTGTCTACAGCTCAGAGATCCGGGACATGTTCCAGCTGCGCTTCTGTTGACAC 485
 QY 481 GACTCCGAGCTCAGGCTGTCTCGAAGATTCTCCGGGACGTGCGCGGAGGGAAGGACTG 540

Db 486 GACTCCGACGTGAGAGTCTTCGAAAGAGTCTCCGGGACGTGCGTCGAGGAGGGACCTG 545
Qy 541 |GAGCAGATTTGAGCGAGTACACACCTTCGGAAGCGGCGCTTCGAGAGTTTCGCTG 600
Db 546 GAGCAGATTTGAGCGAGTACACACCTTCGGAAGCGGCGCTTCGAGAGTTTCGCTG 605
Qy 601 CCGACAAAGAGTATGCGAGTGTATCATCCACAGAGAGTGCACAAATGTTGCCATC 660
Db 606 CCGACAAAGAGTATGCGAGTGTATCATCCACAGAGAGTGCACAAATGTTGCCATC 665
Qy 661 AACCTGATCTGTCAGACATCCAGACATTTGAAATGTCATCTGCAATTCGACCCGA 720
Db 666 AACCTGATCTGTCAGACATCCAGACATTTGAAATGTCATCTGCAATTCGACCCGT 725
Qy 721 GAGAGGTCCAAATGGCGGAGCTACAAAG-CGGACCTTTTCTGAGCGAGGAGCA-CCCTG 778
Db 726 AGAGGCTCAATGGCGGAGCTACAAAGTCCGACCTTTTCTGAGTCCAGGAGCACCCCTG 785
Qy 779 GGATGCTGACCTCTGCGAAACGG 801
Db 786 GGATGCTGACCTCTGCGAAACGG 808

RESULT 7
Bg491358 860 bp mRNA linear EST 27-MAR-2001
LOCUS 602536642F1 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:4684738 5',
DEFINITION mRNA sequence.
ACCESSION Bg491358 GI:13452870
VERSION Bg491358.1 GI:13452870
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 860)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgsab@remail.nih.gov
Tissue Procurement: DCTD/DTF
DNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLU)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLU at:
http://image.llnl.gov
Plate: L10C1493 row: c column: 11
High quality sequence stop: 844.
FEATURES
source 1..860
location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4684738"
/tissue_type="amelanotic melanoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_41"
/note="Organ: skin; Vector: pOTB7; Site: 1: XhoI; Site: 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."

BASE COUNT 195 a 223 c 288 g 154 t
ORIGIN

Query Match 89.5%; Score 746.8; DB 10; Length 860;
Best Local Similarity 99.1%; Fred. No. 3.2e-174;
Matches 803; Conservative 0; Mismatches 2; Indels 5; Gaps 5;

Qy 1 ATGGCTTGCGGAGGCGAAGACTGCGAGGCCCCCGCGGAGGCCGCTCCGAC 60
Db 52 ATGGCTTGCGGAGGCGAAGACTGCGAGGCCCCCGCGGAGGCCGCTCCGAC 111
Qy 61 CAGCGGCCCTTCTGATAGGGGTGAGCGCGGAGCTTCAGCGGGAAGTGCAGTGT 120
Db 112 CAGCGGCCCTTCTGATAGGGGTGAGCGCGGAGCTTCAGCGGGAAGTGCAGTGT 171
Qy 121 GAGAAATCATGAGTGTCTGGGACAGAACAGAGTGGAAACAGGCGGAGGAGTGT 180
Db 172 GAGAAATCATGAGTGTCTGGGACAGAACAGAGTGGAAACAGGCGGAGGAGTGT 221
Qy 181 ATCTGAGCAGAGCAGAGTCTCAAGGTCTTACAGCGCAGAGCAAGGCCCTTG 240
Db 232 ATCTGAGCAGAGCAGAGTCTTACAGGTCTTACAGCGCAGAGCAAGGCCCTTG 291
Qy 241 AAAGCAGTACAAATTTTGACATCCAGATGCTTGTATATGATTTGATGACAGACT 300
Db 292 AAAGCAGTACAAATTTTGACATCCAGATGCTTGTATATGATTTGATGACAGACT 351
Qy 301 CTGAAGAACTGCTGAGAGGCAAAACGCTGAGAGTCCGACCTATGATTTTGACAC 360
Db 352 CTGAAGAACTGCTGAGAGGCAAAACGCTGAGAGTCCGACCTATGATTTTGACAC 411
Qy 361 TCAAGTTTACAGAGACCAAGTGTCTTACCTGCGGACGCTGTTTGAAGGCATC 420
Db 412 TCAAGTTTACAGAGACCAAGTGTCTTACCTGCGGACGCTGTTTGAAGGCATC 471
Qy 421 TTGGTGTCTTACAGCCAGAGATCCGGACATGTTCCACTGCGCTTTTGTGACACC 480
Db 472 TTGGTGTCTTACAGCCAGAGATCCGGACATGTTCCACTGCGCTTTTGTGACACC 531
Qy 481 GACTCCGACGTGAGCGTGTCTGAG-AGTTCCTCCGGAAGTCCGCGAGGAGGACCT 539
Db 532 GACTCCGACGTGAGCGTGTCTGAGAGAGTTCGCGGAGTCCGCGAGGAGGACCT 591
Qy 540 GAGCAGATTTCTGACGACAGTACCAACCTTCGGAAGCGGCTTCGAGAGTTCGCT 599
Db 592 GAGCAGATTTCTGACGACAGTACCAACCTTCGGAAGCGGCTTCGAGAGTTCGCT 651
Qy 600 GCCGACAAAGATATGCGAGTGTATCATCCAGAGAGTGAACATA-TGTTGCCA 658
Db 652 GCCGACAAAGATATGCGAGTGTATCATCCAGAGAGTGAACATA-TGTTGCCA 711
Qy 659 TCAACCTGATCGTGCAGACATCCAGACATCTTGATGTGTGACATCTGAATGGA-C 717
Db 712 TCAACCTGATCGTGCAGACATCCAGACATCTTGATGTGTGACATCTGAATGGA-C 771
Qy 718 CGAGAGGATTCATGAGCGGAGCTTCAAGCGGACCTTTTCTGAGCGAGGACACCT 777
Db 772 CGAGAGGATTCATGAGCGGAGCTTCAAGCGGAGCTTCTGAGCGAGGACACCT 830
Qy 778 GGGATGCTGACCTCT-GGCAACGGTCCACA 806
Db 831 GGGATGCTGACCTCTGGGCAAAAGGTCCACA 860

RESULT 8
Bg826894 916 bp mRNA linear EST 22-MAY-2001
LOCUS 602750978F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:4903819 5',
DEFINITION mRNA sequence.
ACCESSION Bg826894 GI:14174481
VERSION Bg826894.1 GI:14174481
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 916)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)


```

Db      52 ATGGCTTGCGGGGAGGCGAAGACTGCGAGAGCCCCCGCGAGCCGACCGCTCCGAC 111
Qy      61 CAGCGGCCCTTCTGATAGGGGTGAGCGCGGCACTGCGAGGGGAAGTGCACCGTGCT 120
Db      112 CAGGGGCCCTTCTGATAGGGGTGAGCGCGGCACTGCGAGGGGAAGTGCACCGTGCT 171
Qy      121 GAGAAATCATGAGATTGCTGGGACAGAACGAGGTGAAACGCGCAGCGGAGGTGTC 180
Db      172 GAGAAATCATGAGATTGCTGGGACAGAACGAGGTGAAACGCGCAGCGGAGGTGTC 231
Qy      181 ATCTGAGCCGAGACAGATTCTCAAGAGTCTGACGCGAGACGAGAGGCCAAGGCTTG 240
Db      232 ATCTGAGCCGAGACAGATTCTCAAGAGTCTGACGCGAGACGAGAGGCCAAGGCTTG 291
Qy      241 AAAGACAGTCAATTTTGAACCATCCAGATGCGCTTGAATGATTTGATGACAGGACT 300
Db      292 AAAGACAGTCAATTTTGAACCATCCAGATGCGCTTGAATGATTTGATGACAGGACT 351
Qy      301 CTGAAGAACATCGTGAAGGCGAAGAACGCTGAGAGTGCAGACTTATTTGTGACACAC 360
Db      352 CTGAAGAACATCGTGAAGGCGAAGAACGCTGAGAGTGCAGACTTATTTGTGACACAC 411
Qy      361 TCAAGATTACAGAGACCAAGGTGTCTACCTCGGACGCTGCTTCTTTGAGGGCATC 420
Db      412 TCAAGATTACAGAGACCAAGGTGTCTACCTCGGACGCTGCTTCTTTGAGGGCATC 471
Qy      421 TTGGTGTCTACAGCCAGAGATCCGAGGACATGTTCCACTGCGCTCTTGTGAGACAC 480
Db      472 TTGGTGTCTACAGCCAGAGATCCGAGGACATGTTCCACTGCGCTCTTGTGAGACAC 531
Qy      481 GACTCCGAGCTCAGGCTGTCTGAAAGTTCTCCGGACGTCGCGCAGAGGAGGACCTG 540
Db      532 GACTCCGAGCTCAGGCTGTCTGAAAGTTCTCCGGACGTCGCGCAGAGGAGGACCTG 591
Qy      541 GAGAGAGTTCTGAGCGAGATACACACCTTCGTGAAGCCGCGCTTCAGAGAGTTCTGCTG 600
Db      592 GAGAGAGTTCTGAGCGAGATACACACCTTCGTGAAGCCGCGCTTCAGAGAGTTCTGCTG 651
Qy      601 CCGACAAAG-AAATATGCGATGTGATCATCCAGAGAGTGAACAATAT-GGTTGCA 658
Db      652 CCGACAAAGAGTATGCGATGTGATCATCCAGAGAGTGAACAATATGAGTTGCA 711
Qy      659 TCAACTGTATGTGAGACATCAAGACATTTGAAATGTGACATCTGCAATGGCA-C 717
Db      712 TCAACTGTATGTGAGACATCAAGACATTTGAAATGTGACATCTGCAATGGCA-C 771
Qy      718 CGAGAGAGTGTGATGAGCGAGCTACAGAGGACCTTTTCTGAGCCAGAGGACCACTT 777
Db      772 CGAGAGAGTGTGATGAGCGAGCTACAGAGGACCTTTTCTGAGCCAGAGG-ACCACT 829
Qy      778 GGGATGTGACCTCTGGCAACGGTCACTTTG 810
Db      830 GGGATGTGACCTCTGGCAACGGTCACTTTG 862

RESULT 10
LOCUS   AU131406 779 bp mRNA linear EST 01-Aug-2002
DEFINITION AU131406 NT2RP3 Homo sapiens cDNA clone NT2RP3002519 5', mRNA
sequence.
ACCESSION AU131406.1 GI:10991760
VERSION   AU131406
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 779)
Oca,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,
Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and
Isogai,T.
TITLE     HRI human cDNA project
JOURNAL   Unpublished

```

```

COMMENT  Contact: Takao Isogai
          Genomics Laboratory
          Helix Research Institute
          1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
          Tel: 81-438-52-3975
          Fax: 81-438-52-3986
          Email: Genomics@hri.co.jp
          HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
          Research Institute; cDNA library construction: Department of
          Virology, Institute of Medical Science, University of Tokyo, and
          Helix Research Institute.
FEATURES
source
1..779
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="NT2RP3002519"
/cell_type="teratocarcinoma"
/cell_line="NT2"
/clone_1lb="NT2RP3"
/notes="Vector: pME18SFL3; mRNA from NT2 neuronal precursor
cells after 2-weeks retinoic acid (RA) induction"
BASE COUNT  179 a 197 c 253 g 147 t 3 others
ORIGIN
Query Match 87.1%; Score 726.8; DB 9; Length 779;
Best Local Similarity 99.1%; Pred. No. 2,7e-169;
Matches 750; Conservative 0; Mismatches 5; Indels 2; Gaps 2;
Qy      1 ATGGCTTGCGGGGAGGCGAAGACTGCGAGAGCCCGCGCGAGCCGACCGCTCCGAC 60
Db      24 ATGGCTTGCGGGGAGGCGAAGACTGCGAGAGCCCGCGCGAGCCGACCGCTCCGAC 83
Qy      61 CAGGGGCCCTTCTGATAGGGGTGAGCGCGGCACTGCGAGGGGAAGTGCACCGTGCT 120
Db      84 CAGGGGCCCTTCTGATAGGGGTGAGCGCGGCACTGCGAGGGGAAGTGCACCGTGCT 143
Qy      121 GAGAAATCATGAGATTGCTGGGACAGAACGAGGTGAAACGCGGACGCGAAGGTGTC 180
Db      144 GAGAAATCATGAGATTGCTGGGACAGAACGAGGTGAAACGCGGACGCGAAGGTGTC 203
Qy      181 ATCTGAGCCGAGACAGATTCTCAAGAGTCTGACGCGCAGAGACGAAGGCCCTTG 240
Db      204 ATCTGAGCCGAGACAGATTCTCAAGAGTCTGACGCGCAGAGACGAAGGCCCTTG 263
Qy      241 AAAGACAGTCAATTTTGAACCATCCAGATGCGCTTGAATGATTTGATGACAGGACT 300
Db      264 AAAGACAGTCAATTTTGAACCATCCAGATGCGCTTGAATGATTTGATGACAGGACT 323
Qy      301 CTGAAGAACATCGTGAAGGCGAAGAACGCTGAGAGTCCGACCTATGATTTTGTGACACAC 360
Db      324 CTGAAGAACATCGTGAAGGCGAAGAACGCTGAGAGTCCGACCTATGATTTTGTGACACAC 383
Qy      361 TCAAGATTACAGAGACCAAGGTGTCTACCTCGGACGCTGCTTCTTTGAGGGCATC 420
Db      384 TCAAGATTACAGAGACCAAGGTGTCTACCTCGGACGCTGCTTCTTTGAGGGCATC 443
Qy      421 TTGGTGTCTACAGCCAGAGATCCGAGGACATGTCACCTGCGCTCTTGTGAGACAC 480
Db      444 TTGGTGTCTACAGCCAGAGATCCGAGGACATGTCACCTGCGCTCTTGTGAGACAC 503
Qy      481 GACTCCGAGCTCAGGCTGTCTGAAAGATTCTCCGAGACGTCGCGCAGAGGAGGACCTG 540
Db      504 GACTCCGAGCTCAGGCTGTCTGAAAGATTCTCCGAGACGTCGCGCAGAGGAGGACCTG 563
Qy      541 GAGAGATTCTGACGAGTACACACCTTCGTGAAGCCGCGCTTCTGAGAGATTCTGCTG 600
Db      564 GAGAGATTCTGACGAGTACACACCTTCGTGAAGCCGCGCTTCTGAGAGATTCTGCTG 623
Qy      601 CCGACAAAGAGATGCGAGATGATCATCCACGAGAGGTGGAACAATATGTTGCGATC 660
Db      624 CCGACAAAGAGATGCGAGATGATCATCCACGAGAGGTGGAACAATATGTTG-CAATC 682

```

Qy 661 AACCTGATCGTGAGACATCCAGCATTTCTGATGTGATCATCTCAATGAGCAGCGA 720
 Db 683 AACCTGATCGTGAGACATCCAGCATTTCTGATGTGATCATCTCAATGAGCAGCGA 741
 Qy 721 GGAAGGTTCATGGCGGAGCTCAAGACGGACCTTTT 757
 Db 742 GGANGGTCAATGGCGGAGCTCAAGACGGACCTTTT 778

RESULT 11
 BM917506
 LOCUS 1044 bp mRNA linear EST 12-MAR-2002
 DEFINITION AGENCOURT 6605568 NIH_MGC_106 Homo sapiens cDNA clone IMAGE:5484033
 5', mRNA sequence.
 ACCESSION BM917506
 VERSION BM917506.1 GI:19367885
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 1044)
 AUTHORS NIH-MGC http://mgs.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabs-remail.nih.gov
 Tissue Procurement: Dr. Daniel McVicar, DBS/NCI
 cDNA Library Preparation: Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLCM2012 row: k column: 10
 High quality sequence stop: 624.
 Location/Qualifiers
 1..1044
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5484033"
 /issue_type="natural killer cells, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_id="NIH_MGC_106"
 /note="Organ: blood; Vector: pOTB7; Site_1: XhoI; Site_2:
 EcoRI; cDNA made by oligo-dT priming. Directionally cloned
 into EcoRI/XhoI sites using the following 5' adaptor:
 GGACACGAG(G). Library constructed by Ling Hong in the
 laboratory of Gerald M. Rubin (University of California,
 Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
 Superscript II RT (Life Technologies). Note: this is a
 NIH_MGC Library."

BASE COUNT 224 a 306 c 318 g 196 t
 ORIGIN

Query Match 86.9%; Score 724.8; DB 12; Length 1044;
 Best Local Similarity 98.9%; Pred. No. 9,7e-169;
 Matches 740; Conservative 0; Mismatches 7; Indels 1; Gaps 1;

Qy 13 GGAAGGCAAGACTGCGAGAGCCCGGCGGAGGCGGACCGTCCGACCGAGCGGCCCTTC 72
 Db 1 GGAAGGCAAGACTGCGAGAGCCCGGCGGAGGCGGACCGTCCGACCGAGCGGCCCTTC 60
 Qy 73 CTGATAGGGGTGAGCGCGGCACTGCCAGCGGAAAGTGAACCGTGTGTGAGAAAGATCATG 132
 Db 61 CTGATAGGGGTGAGCGCGGCACTGCCAGCGGAAAGTGAACCGTGTGTGAGAAAGATCATG 120
 Qy 133 GAGTGTCTGGAGCAAGACGAGGTGGAACAGCGGGAAGGTGTGATCTTGAGCCAG 192
 Db 121 GAGTGTCTGGAGCAAGACGAGGTGGAACAGCGGGAAGGTGTGATCTTGAGCCAG 180
 Qy 193 GACAGGTTTACAAGGTCCTGACGGCAGAGCAAGAGCCAAAGGCTTGAAAGGACAGTAC 252

Db 181 GACAGGTTTACAAGGTCCTGACGGCAGAGCAGAAAGCCAAAGGCTTGAAAGGACAGTAC 240
 Qy 253 AATTGTTGACCTTCAGATGCTTTGATATGATTTATGACAGGACTCTGAAAGAACATC 312
 Db 241 AATTGTTGACCTTCAGATGCTTTGATATGATTTATGACAGGACTCTGAAAGAACATC 300
 Qy 313 GTGAGGAGGCAAAAGGTTGAGGAGGCGGACCTATGATTTGTGACACACTCAAGTTACCA 372
 Db 301 GTGAGGAGGCAAAAGGTTGAGGAGGCGGACCTATGATTTGTGACACACTCAAGTTACCA 360
 Qy 373 GAGACCAAGGTGCTTACCTCGGAGCGTGTGTTGAGGAGCATCTTGTTCTAC 432
 Db 361 GAGACCAAGGTGCTTACCTCGGAGCGTGTGTTGAGGAGCATCTTGTTCTAC 420
 Qy 433 AGCCAGAGATCCGGGACATGTTCCACTGCGCTTTGTGTGACACCGACTCCAGCTC 492
 Db 421 AGCCAGAGATCCGGGACATGTTCCACTGCGCTTTGTGTGACACCGACTCCAGCTC 480
 Qy 493 AGGCTGTCTCGAAGATTCGCGGAGCGTGGCGCGAGGAGGACCTGGAGCATTCCTG 552
 Db 481 AGGCTGTCTCGAAGATTCGCGGAGCGTGGCGCGAGGAGGACCTGGAGCATTCCTG 540
 Qy 553 AGCAGTACACCACTTGTGAAAGCCGCTTCGAGAGTTCCTGCGCAGCAAGAAAG 612
 Db 541 AGCAGTACACCACTTGTGAAAGCCGCTTCGAGAGTTCCTGCGCAGCAAGAAAG 600
 Qy 613 TATGCCGATGTATCATCCACGAGAGTGAGACATATGTTGCCATCAACTGATCTG 672
 Db 601 TATGCCGATGTATCATCCACGAGAGTGAGACATATGTTGCCATCAACTGATCTG 660
 Qy 673 CAGCATATCAGGACATTCGTAATGTGTGATCATCTGAAATGGACGAGAGGTCAT 732
 Db 661 CAGCATATCAGGACATTCGTAATGTGTGATCATCTGAAATGGACGAGAGGTCAT 720
 Qy 733 GGAG-AGAGCTACAGCGGACCTTTTCT 759
 Db 721 GGCGGGAGCTACAGCGGACCTTTT 748

RESULT 12
 B0935919 922 bp mRNA linear EST 21-AUG-2002
 LOCUS B0935919
 DEFINITION AGENCOURT 6785353 NIH_MGC_18 Homo sapiens cDNA clone IMAGE:6371179
 5', mRNA sequence.
 ACCESSION B0935919
 VERSION B0935919.1 GI:22351302
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 922)
 AUTHORS NIH-MGC http://mgs.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabs-remail.nih.gov
 Tissue Procurement: DCTD/DRP/Gazdar
 cDNA Library Preparation: Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLCM2544 row: o column: 20
 High quality sequence stop: 625.
 Location/Qualifiers
 1..922
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6371179"

/tissue_type="large cell carcinoma"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_18"
 /note="Organ: lung; Vector: pOB7; Site 1: XhoI; Site 2:
 EcoRI; cDNA made by oligo-dt priming. Directionally cloned
 into EcoRI/XhoI sites using the following 5' adaptor:
 GGCAAGGAG(G). Library constructed by Ling Hong in the
 laboratory of Gerald M. Rubin (University of California,
 Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
 Superscript II RT (Life Technologies). Note: this is a
 NIH_MGC Library."

BASE COUNT 203 a 255 c 297 g 161 t 6 others

Query Match 85.6%; Score 714.2; DB 13; Length 922;
 Best Local Similarity 96.1%; Pred. No. 3.9e-165;
 Matches 805; Conservative 0; Mismatches 26; Indels 7; Gaps 7;

1 ATGGCTTCGGCGGAGGCGAAGACTGCGAGAGCCCGCGCGGAGCCGACCGTCCGAC 60
 36 ATGGCTTCGGCGGAGGCGAAGACTGCGAGAGCCCGCGCGGAGCCGACCGTCCGAC 95
 61 CAGCGGCCCTTCCTGATGAGGCGGCGGCACTGCGAGCGGAGTGCAGCTGTGT 120
 96 CAGCGGCCCTTCCTGATGAGGCGGCGGCACTGCGAGCGGAGTGCAGCTGTGT 155
 121 GAGAAGATCATGAGTTCGCGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 180
 156 GAGAGATCATGAGTTCGCGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 215
 181 ATCTGAGCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 240
 216 ATCTGAGCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 275
 241 AAAAGACAGTCAATTTTGAACATCCAGATGCTTGTATATGATTTGATGACAGACT 300
 276 AAAAGACAGTCAATTTTGAACATCCAGATGCTTGTATATGATTTGATGACAGACT 335
 301 CTGAAGAACATGTCGAGAGGCGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 360
 336 CTGAAGAACATGTCGAGAGGCGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 395
 361 TCAAGGTTACGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 420
 396 TCAAGGTTACGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 455
 421 TTGGTGTCTACAGCGAGAGATCCGAGACATGTTCCACTGCGCTCTTGTGTAACAC 480
 456 TTGGTGTCTACAGCGAGAGATCCGAGACATGTTCCACTGCGCTCTTGTGTAACAC 515
 481 GACTCCGAGCTCAGAGCTGTCTCGAAGAGTCTCCGAGGAGGAGGAGGAGGAGGAGG 540
 516 GACTCCGAGCTCAGAGCTGTCTCGAAGAGTCTCCGAGGAGGAGGAGGAGGAGGAGG 575
 541 GAGCAGATTCTGAGCGAGTACACCACTTGTGAAAGCGGCGCTTCAAGAGTTCGCTG 600
 576 GAGCAGATTCTGAGCGAGTACACCACTTGTGAAAGCGGCGCTTCAAGAGTTCGCTG 635
 601 CCGCAAGAAAGATGTC-CCATGATGAT-CCCAAGAGAGTGG-ACATATGTTGGC 657
 636 CCGCAAGAAAGATGTCGATGATGATCCCAAGAGAGTGG-ACATATGTTGGC 695
 658 ATCAACCTGATCGAGCAGCATCAG-GACATTCTGAATGAGTGAATGCAATGGA 716
 696 ATCAACCTGATCGAGCAGCATCAGGACATCTGATATGTAACATCTGCAATGGA 755
 717 CCGA-GGAGGTCGAATGAGGCGAGCTACAAGCGGAGCTTTCTGAGCGAGGAGGAG-CA 774
 756 CCGAGGAGGTCGAATGAGGCGAGCTACAAGCGGAGCTTTCTGAGAGGAGGAGGAGG 815
 775 CTTGGAGTCTGACCTCTGGCA-ACGGTCACTTTGAGTTCAGCAGCAGAGCCCAAC 831
 816 CTTGGAGTCTGACCTCTGGCAACGNNCCCATTTTGGAGTCAAGCAAAACCCCCC 873

RESULT 13
 BI258532 848 bp mRNA linear EST 17-JUL-2001
 LOCUS 602972340F1 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:5112032 5',
 DEFINITION mRNA sequence.
 ACCESSION BI258532
 VERSION BI258532.1 GI:14814971
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE NIH-MGC http://mgc.nci.nih.gov/
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 TITLE Unpublished
 JOURNAL Contact: Robert Strausberg, Ph.D.
 COMMENT Email: cgabs-remail.nih.gov
 Tissue Procurement: ATCC

CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: Incyte Genomics, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LHM11271 row: 0 column: 09
 High quality sequence stop: 839.

FEATURES
 source Location/Qualifiers
 1..848
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5112032"
 /tissue_type="cervical carcinoma cell line"
 /lab_host="DH10B"
 /note="Organ: cervix; Vector: pCMV-SPORT6; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally. Primer: 'oligo dt.
 Average insert size 1.4 kb. Library prepared by Life
 Technologies."

BASE COUNT 196 a 219 c 271 g 162 t

Query Match 85.2%; Score 710.2; DB 12; Length 848;
 Best Local Similarity 95.2%; Pred. No. 3.6e-165;
 Matches 797; Conservative 0; Mismatches 8; Indels 32; Gaps 5;

1 ATGGCTTCGGCGGAGGCGAAGACTGCGAGAGCCCGCGCGGAGCCGACCGTCCGAC 60
 6 ATGGCTTCGGCGGAGGCGAAGACTGCGAGAGCCCGCGCGGAGCCGACCGTCCGAC 65
 61 CAGCGGCCCTTCCTGATGAGGCGGCGGCACTGCGAGCGGAGTGCAGCTGTGT 120
 66 CAGCGGCCCTTCCTGATGAGGCGGCGGCACTGCGAGCGGAGTGCAGCTGTGT 98
 121 GAGAAGATCATGAGTTCGCGGAGAGAGAGAGTGAACAGGCGGAGGAGGAGGAGGAGG 180
 99 GAGAAGATCATGAGTTCGCGGAGAGAGAGAGTGAACAGGCGGAGGAGGAGGAGGAGG 158
 181 ATCTGAGCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 240
 159 ATCTGAGCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 218
 241 AAAAGACAGTCAATTTTGAACATCCAGATGCTTGTATATGATTTGATGACAGACT 300
 219 AAAAGACAGTCAATTTTGAACATCCAGATGCTTGTATATGATTTGATGACAGACT 278
 301 CTGAAGAACATGTCGAGAGGCGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 360
 279 CTGAAGAACATGTCGAGAGGCGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 338

Oy	361	CAAGGTTACAGAGACCAAGCGTGTCTAACCTCGGGAGCGTGGTCTGTGTTGAGGGCAATC	420
Db	339	TCAGGTTACAGAGACCAAGGTGTCTAACCTCGGGAGCGTGGTCTGTGTTGAGGGCAATC	398
Oy	421	TTGGTGTCTACAGCCAGAGATCCGGGACATGTTCCACTTGCGGCTCTTCGTGACACC	480
Db	399	TTGGTGTCTACAGCCAGAGATCCGGGACATGTTCCACTTGCGGCTCTTCGTGACACC	458
Oy	481	GATTCGCGACGTCAAGGCTGTCTTGAAAGATTCTTCGGGACGTGCGCCGAGAGAGGACCTG	540
Db	459	GATTCGCGACGTCAAGGCTGTCTTGAAAGATTCTTCGGGACGTGCGCCGAGAGAGGACCTG	518
Oy	541	GAGCAGATTCTGACCCAGCAGTACACCACTTCGTGAAGCGGACCTTCGAGAGATTCTGCCTG	600
Db	519	GAGCAGATTCTGACCCAGCAGTACACCACTTCGTGAAGCGGACCTTCGAGAGATTCTGCCTG	578
Oy	601	CCGACAAAGAAAGTATGCCGATGTGATCATCCACAGGAGTGGACAAATATGTTGCCATC	660
Db	579	CCGACAAAGAAAGTATGCCGATGTGATCATCCACAGGAGTGGACAAATATGTTGCCATC	638
Oy	661	AACCTGATCGAGCAGCATCCAGGACATTTCTGAACTGTGTACATCTGCAATGTGGCACCG	719
Db	639	AACCTGATCGAGCAGCATCCAGGACATTTCTGAACTGTGTACATCTGCAATGTGGCACCG	698
Oy	720	AGAGGGGTCCAATGGGCGG--AGCTACAAAGCGGACCTTTTCTGAGCCAGGGGACCAACCT	777
Db	699	AGAGGGGTCCAATGGGCGGAGCTTCAAGGGGACCTTTTCTGAGCCAGGGGACCAACCT	758
Oy	778	GGGATGCTGACCTCTGCGCAAC-GGTCACTTTGG-AGTCCAGCAGCAGACCCCACT	832
Db	759	GGGATGCTGACCTCTGCGCAACGGGTCACTTTGGAAATCCAGCAGCAAGACCCCAAT	815

RESULT 14	LOCUS	DEFINITION
B0072501	B0072501	1036 bp mRNA
	AGENCOURT 6838909	linear EST 02-APR-2002
	NIH_MGC_122	Homio sapiens CDNA clone IMAGE:5761699
	5', mRNA sequence.	

VERSION	BQ072501.1	GI:19901547
KEYWORDS	Est.	
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	
	Eukaryotes; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
REFERENCE	1 (bases 1 to 1036)	
AUTHORS	NH-MGC ftp://mgc.nci.nih.gov/.	
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)	
JOURNAL	Unpublished	
COMMENT	Contact: Robert Strausberg, Ph.D.	

FEATURES

```

1. .1036="Homo sapiens"
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5761694"
/lab_host="DH10B"
/clone_lib="NH MGC 122"
/note="Organ: pooled lung and spleen; Vector: pCMV-Sport6,
Site 1: NotI; Site 2: EcoRV (destroyed); RNA source
anonymous pool of 52 week female lung, 16 week female
spleen, and 20-22 week male spleen. Library is oligo-dT
primed and directionally cloned (EcoRV site is destroyed

```

upon cloning). Average insert size 1.4 kb, insert size range 1-3 kb. Library is normalized and knibbed for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 026. Note: this is a NIH MOC Library."

Query Match	85.0%;	Score 709;	DB 13;	Length 1036;
Best Local Similarity	97.2%;	Pred. No. 7.9e-165;		
Matches 765;	Conservative	0;	Mismatches 15;	Indels 7;
				Gaps 4;

QY	1	ATGGCTTCGGCGGAGGCGAAGACTGCGAGAGCGCCGCGCGGAGGCGAGCCGTCCGAC	60
Db	64	ATGCTTCGGCGGGAGGCGAAGACTGCGAGAGCCCGCGCGGAGGCGAGCCGTCCGAC	123
QY	61	CAGCGGCCCTTCTCTGATAGGGGTGAGCGGCGGCACTGCCAGCGGGAGTGCACCTGTGT	120
Db	124	CAGCGGCCCTTCTCTGATAGGGGTGAGCGGCGGCACTGCCAGCGGGAGTGCACCTGTGT	183
QY	121	GAGAAATCATGGAGTTGCTGGGGAACAAGAGGTGGAACAAGCGGCGCGGAAGTGGTC	180
Db	184	GAGAAATCATGGAGTTGCTGGGGAACAAGAGGTGGAACAAGCGGCGCGGAAGTGGTC	243
QY	181	ATCCTGAGCAGAGACAGTTCTTCAAGAGTCTGTACGGCAAGACAAGAGCAGAGCCTTG	240
Db	244	ATCCTGAGCAGAGACAGTTCTTCAAGAGTCTGTACGGCAAGACAAGAGCAGAGCCTTG	303
QY	241	AAAGGACAGTACAATTTTGACATCCAGATGCTCTTGATATATGATTTGATGCACAGACT	300
Db	304	AAAGGACAGTACAATTTTGACATCCAGATGCTCTTGATATGATTTGATGCACAGACT	363
QY	301	CTGAAATAATCGTGGAGGGCAAAACGTGGAGTGGCGGACATATGATTTTGTGACACAC	360
Db	364	CTGAAATAATCGTGGAGGGCAAAACGTGGAGTGGCGGACATATGATTTTGTGACACAC	423
QY	361	TCAGGTTTACAGAGACCAACGAGTGTCTACCTCGCGGACGTGGTCTGTTTGAAGGATC	420
Db	424	TCAGGTTTACAGAGACCAACGAGTGTCTACCTCGCGGACGTGGTCTGTTTGAAGGATC	483
QY	421	TTGGTGTCTACAGCCAGAGAGATCCGGGACATGTTTCCACTCGCGCTTCTGTGGAACAC	480
Db	484	TTGGTGTCTACAGCCAGAGAGATCCGGGACATGTTTCCACTCGCGCTTCTGTGGAACAC	543
QY	481	GACTCCGAGTCAAGGCTGTCTGAAAGATCTCCGGGACGTGGGCGGAGGGAGGACCTG	540
Db	544	GACTCCGAGTCAAGGCTGTCTGAAAGATCTCCGGGACGTGGGCGGAGGGAGGAGCCTG	603
QY	541	GAGCAATTTCTGACGCGAGTACACACCTTCTGTAAAGCGGCTTTCAGAGATTTGCTCG	600
Db	604	GAGCAATTTCTGACGCGAGTACACACCTTCTGTAAAGCGGCTTTCAGAGATTTGCTCG	663
QY	601	CCGACAAAGAGTATGCCGATGTGATCATCCACAGAGAGTGGACAAATATGTTGCCAT	659
Db	664	CCGACAAAGAGTATGCCGATGTGATCATCCACAGAGAGTGGACAAATATGTTGCCAT	723
QY	660	CAACCTGATCTGCGACACATCCAGGACATTTGAAATGATGACCTTGTGGACAC	718
Db	724	CAACCTGATCTGCGACACATCCAGGACATTTGAAATGATGACCTTGTGGACAC	783
QY	719	GAGG--AGGATCCAAATGGGCGAGTACACAAGCGGACCTTTTCTGAG--CCAGGGAGCA	773
Db	784	GAGGAGGGGTCAACATGGGCGAGCTACAAAGCGGACCTTTTCTGAGCCAGGGAGCA	843
QY	774	CCCTGGG 780	
Db	844	CCCTGG 850	

RESULT 15	BI261258	LOCUS	827 bp	mRNA	linear	EST 17-Jul-2001
DEFINITION	BI261258	602969110F1 NIH_MGC_12	Homo sapiens cDNA clone IMAGE:5108666 5',			

mRNA sequence.
 ACCESSION B1261258
 VERSION B1261258.1 GI:14820349
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 827)
 NIH-MGC http://mgi.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 AUTHORS Unpublished
 JOURNAL Contact: Robert Strausberg, Ph.D.
 COMMENT Email: cga@bbs-remail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: Incyte Genomics, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/MLNL at:
 http://image.llnl.gov
 Plate: LHAM1263 Row: C Column: 03
 High quality sequence start: 5
 High quality sequence stop: 799.
 Location/Qualifiers
 1..827
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5108666"
 /tissue_type="cervical carcinoma cell line"
 /lab_host="DH10B"
 /clone_lib="NIH MGC 12"
 /note="Organ: cervix; Vector: pCMV-Sport6; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
 Average insert size 1.4 Kb. Library prepared by Life
 Technologies."
 BASE COUNT 189 a 210 c 274 g 154 t
 ORIGIN
 Query Match 84.9%; Score 708.2; DB 12; Length 827;
 Best Local Similarity 97.2%; Pred. No. 1.1e-164;
 Matches 753; Conservative 0; Mismatches 18; Indels 4; Gaps 3;
 QY 1 ATGGCTTCGGGCGGAGGCGAAGACTGCGAGAGCCCGCGGAGCGGACCGTCCGCAC 60
 DB 52 ATGGCTTCGGGCGGAGGCGAAGACTGCGAGAGCCCGCGGAGCGGACCGTCCGCAC 111
 QY 61 CAGCGGCCCTTCTGATAGGGGTGAGCGCGGCACTGCGAGCGGGAAGTCGACCGTGTGT 120
 DB 112 CAGCGGCCCTTCTGATAGGGGTGAGCGCGGCACTGCGAGCGGGAAGTCGACCGTGTGT 171
 QY 121 GAGAGATCATGAGTGTCTGGGACAGACGAGGTGGAACAGCGGCAAGCGGAGGTGTC 180
 DB 172 GAGAGATCATGAGTGTCTGGGACAGACGAGGTGGAACAGCGGCAAGCGGAGGTGTC 231
 QY 181 ATCTGAGCCGAGCAGAGTCTCAAGGTCTGACGGGAGGAGGAGGCGGAGGCTTG 240
 DB 232 ATCTGAGCCGAGCAGAGTCTCAAGGTCTGACGGGAGGAGGAGGCGGAGGCTTG 291
 QY 241 AAGAGCAGTACATTTTGAATCCATCCAGATGCTTTGATATATGATTGATGACAGGACT 300
 DB 292 AAGAGCAGTACATTTTGAATCCATCCAGATGCTTTGATATATGATTGATGACAGGACT 351
 QY 301 CTGAAGAACATCTGAGAGGGGCAAAAGCGTGAAGTCCGACCTATGATTTTGTGACAC 360
 DB 352 CTGAAGAACATCTGAGAGGGGCAAAAGCGTGAAGTCCGACCTATGATTTTGTGACAC 411
 QY 361 TCAAGTTTACAGAGACCAAGGTGTCTACCTCGGAGAGCGGTCTGTTGAGGGGATC 420
 DB 412 TCAAGTTTACAGAGACCAAGGTGTCTACCTCGGAGAGCGGTCTGTTGAGGGGATC 471
 QY 421 TTGGTGTCTACAGCCAGAGATCCGGAGCATGTTCCACCTGCGCTCTTGTGACACC 480

DB 472 TTGGTGTCTACAGCCAGAGATCCGGAGCATGTTCCACCTGCGCTCTTGTGACACC 531
 QY 481 GACTCCGACGTCAAGGTGTCTCGAAGATTCT-CCGGAGCTGCGCGAGGAGGACCT 539
 DB 532 GACTCCGACGTCAAGGTGTCTCGAAGATTCTCCCGGAGCTGCGCGAGGAGGACCT 591
 QY 540 GAGGAGATTCTGACGAGTACACCACTTCGGAAGCGGCGCTTGAGAGTTCGCT 599
 DB 592 GAGGAGATTCTGACGAGTACACCACTTCGGAAGCGGCGCTTGAGAGTTCGCT 651
 QY 600 GCCGACAAAGATATGCGGATGTGATCCACAGAGAGTGAACATATGTTGACAT 659
 DB 652 GCCGACAAAGATATGCGGATGTGATCCACAGAGAGTGAACATATGTTGACAT 711
 QY 660 CAACCTGATGTGACAGACAT--CCAGGACATTCTGAT-GTGAACATCTGCAATGCA 716
 DB 712 CAACCTGATGTGACAGACATCTCCAGGACATTTGATGAGTGTGACATCTGCAATGCA 771
 QY 717 CCGAGGAGGTTCGATGAGGCGGAGCTACAAAGCGGACCTTTCTGAGCGAGGGAC 771
 DB 772 CCGAGGAGGTTCGATGAGGCGGAGCTACAAAGCGGACCTTTCTGAGCGAGGGAC 826

Search completed: November 25, 2003, 02:00:51
 Job time : 2012.96 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 25, 2003, 02:01:02 ; Search time 2006.96 Seconds
(without alignments)
10099.814 Million cell updates/sec.

Title: US-09-896-522-3

Perfect score: 834

Sequence: 1 atgctctgcgcgagagcgca.....ccagcagcagaccacactga 834

Scoring table: OLIGO_NUC

Searched: 22781392 seqs, 12152238056 residues

Word size : 15

Total number of hits satisfying chosen parameters: 22821

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-Processing: Listing first 45 summaries

Database :

```
.EST:*
1:  em_estba:*
2:  em_estbm:*
3:  em_estin:*
4:  em_estnu:*
5:  em_estov:*
6:  em_estpl:*
7:  em_estro:*
8:  em_hic:*
9:  gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gse_hum:*
18: em_gse_inv:*
19: em_gse_pln:*
20: em_gse_vrt:*
21: em_gse_fun:*
22: em_gse_mam:*
23: em_gse_mus:*
24: em_gse_pro:*
25: em_gse_rod:*
26: em_gse_phg:*
27: em_gse_vrt1:*
28: gb_gse1:*
29: gb_gse2:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	708	84.9	1044	12	BM917506 AGENCOURT
2	690	82.7	1118	12	BM467984 AGENCOURT
3	684	82.0	898	14	CD514811 AGENCOURT
4	677	81.2	725	12	BI256928 BI256928 602975534

5	650	77.9	916	12	BG826894	BG826894 602750978
6	643	77.1	1201	13	BX394295	BX394295 BX394295
7	640	76.7	730	12	BG770518	BG770518 602734236
8	619	74.2	956	11	BQ719741	BQ719741 AGENCOURT
9	615	73.7	989	10	BG390519	BG390519 602416183
10	611	73.3	869	10	BG491384	BG491384 60235670
11	605	72.5	779	9	AU131406	AU131406 AU131406
12	602	72.2	1036	9	BQ072501	BQ072501 AGENCOURT
13	589	70.6	848	12	BI258532	BI258532 602972340
14	582	69.8	1055	12	BM545603	BM545603 AGENCOURT
15	580	69.5	661	12	BI830878	BI830878 603081002
16	577	69.2	922	13	BQ935919	BQ935919 AGENCOURT
17	575	68.9	753	9	AU122008	AU122008 AU122008
18	575	68.9	910	13	BQ437115	BQ437115 AGENCOURT
19	558	66.9	959	12	BI770572	BI770572 60360368
20	557	66.8	963	14	CD359372	CD359372 AGENCOURT
21	556	66.7	796	9	AU141725	AU141725 AU141725
22	535	64.1	827	12	BI261258	BI261258 602699110
23	535	64.1	998	13	BX400889	BX400889 K-EST0091
24	533	63.9	554	12	BM822058	BM822058 K-EST0091
25	531	63.7	778	14	CD101693	CD101693 AGENCOURT
26	530	63.5	860	10	BG491358	BG491358 60235642
27	525	62.9	584	10	AW961647	AW961647 EST373720
28	520	62.4	1114	13	BX343101	BX343101 BX343101
29	508	60.9	799	10	BG723578	BG723578 602694293
30	503	60.3	1039	12	BM803129	BM803129 AGENCOURT
31	484	59.2	810	12	BI753007	BI753007 603025788
32	491	58.9	514	12	BM827668	BM827668 K-EST0100
33	491	58.9	515	12	BM835278	BM835278 K-EST0110
34	480	57.6	923	10	BE891861	BE891861 601435415
35	480	57.6	1095	12	BM802939	BM802939 AGENCOURT
36	464	55.6	797	12	BI116432	BI116432 602688854
37	458	54.9	1097	10	BF347745	BF347745 602022586
38	451	54.1	819	12	BI910015	BI910015 603068126
39	429	51.4	585	14	CB156232	CB156232 K-EST0214
40	424	50.8	945	13	BQ677293	BQ677293 AGENCOURT
41	420	50.4	883	13	BUB38101	BUB38101 AGENCOURT
42	420	50.4	894	10	BG334621	BG334621 602460907
43	420	50.4	948	13	BO680950	BO680950 AGENCOURT
44	420	50.4	973	13	BO683188	BO683188 AGENCOURT
45	419	50.2	1219	12	BM545221	BM545221 AGENCOURT

ALIGNMENTS

RESULT 1
BM917506 1044 bp mRNA linear EST 12-MAR-2002
LOCUS AGENCOURT_6606568 NIH_MGC_106 Homo sapiens cDNA clone IMAGE:5484033
DEFINITION 5', mRNA sequence.
ACCESSION BM917506.1 GI:19367885
VERSION BM917506
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-r@mail.nih.gov
Tissue Procurement: Dr. Daniel McVicar, DBS/NCI
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LINC2012 row: k column: 10
High quality sequence stop: 624.

QY 421 TTGGTGTCTACAGCCAGAGATCCGGGACATGTTCCACCTGGGCCCTTCTGTGGACACC 480
|
|
Db 467 TTGGTGTCTACAGCCAGAGATCCGGGACATGTTCCACCTGGGCCCTTCTGTGGACACC 526
|
|
QY 481 GACTCCGACGTGAGCTGTCTCGAAGAGTTCTCCGGACGTGGCCGAGGAGGACCTG 540
|
|
Db 527 GACTCCGACGTGAGCTGTCTCGAAGAGTTCTCCGGACGTGGCCGAGGAGGACCTG 586
|
|
QY 541 GAGCAGATTCTGACGACGATACACCACTTGTGTGAAGCCGCTTCGAGAGTTCTGCTG 600
|
|
Db 587 GAGCAGATTCTGACGACGATACACCACTTGTGTGAAGCCGCTTCGAGAGTTCTGCTG 646
|
|
QY 601 CCGACAAAGAGATGCGGATGTGATCATCCACGAGAGATGACAAATATGTTGGCATC 660
|
|
Db 647 CCGACAAAGAGATGCGGATGTGATCATCCACGAGAGATGACAAATATGTTGGCATC 706
|
|
QY 661 AACCTGATCGTGACGACATCCAGACATT 690
|
|
Db 707 AACCTGATCGTGACGACATCCAGACATT 736
|
|
RESULT 3 898 bp mRNA linear EST 06-JUN-2003
CD514811
LOCUS IMAGE:30356201 5', mRNA sequence.
DEFINITION AGENCOURT 14376196 NIH_MGC_181 Homo sapiens CDNA clone
CD514811
VERSION CD514811.1 GI:31446529
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 898)
NIH-MGC http://mgi.nci.nih.gov/.
TITLES National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgabbs-remail.nih.gov
Tissue Procurement: Dr. Michael Brownstein
CDNA Library Preparation: Invitrogen Corp
DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: NDAM477 row: 1 column: 10
High quality sequence stop: 662.
Location/Qualifiers
1. 898
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30356201"
/tissue_type="White Matter"
/dev_stage="Unknown"
/lab_host="DH10B-Ton A (T1 and T5 phage resistances)"
/note="Vector: pCMV-Sport6.1; Site_1: NotI; Site_2: EcoRV
(destroyed); Library is oligo-dT primed and directionally
cloned (EcoRV site is destroyed upon cloning). Average
insert size 1.42 kb. Library was constructed by
(Invitrogen). Note: this is a NIH MGC Library."
BASE COUNT 194 a 231 c 310 g 158 t 5 others
ORIGIN
Query Match 82.0%; Score 684; DB 14; Length 898;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 684; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCTTCGGCGGAGAGGAGAACTGCGAGAGCCCGCGCGGAGGCGGACCGTCCGAC 60
|
|
Db 72 ATGGCTTCGGCGGAGAGGAGAACTGCGAGAGCCCGCGCGGAGGCGGACCGTCCGAC 131
|
|
QY 61 CAGCGGCCCTTCCTGATTAAGGGGTGAGCGCGGCACTGCGAGGAGAGTGCACCTGTGT 120
|
|
Db 132 CAGCGGCCCTTCCTGATTAAGGGGTGAGCGCGGCACTGCGAGGAGAGTGCACCTGTGT 191
|
|
QY 121 GAGAGATCATGAGAGTTGCTGGGACAGAAAGAGTGAACAGCGGACGGAAGTGTGTC 180
|
|
Db 192 GAGAGATCATGAGAGTTGCTGGGACAGAAAGAGTGAACAGCGGACGGAAGTGTGTC 251
|
|
QY 181 ATCTTAAGCCAGAGACAGGTCTTCAAGGTCTGACGCGGACGACAGAAAGCCGCTTG 240
|
|
Db 252 ATCTTAAGCCAGAGACAGGTCTTCAAGGTCTGACGCGGACGACAGAAAGCCGCTTG 311
|
|
QY 241 AAAGCAGTACAAATTTTGAACCATCCAGATGCTTGTATATGATTTGATGACAGACT 300
|
|
Db 312 AAAGCAGTACAAATTTTGAACCATCCAGATGCTTGTATATGATTTGATGACAGACT 371
|
|
QY 301 CTGAGAACATCGTGAGGCGCAAAACGATGAGGTGCGGACCTATGATTTTGTGACAC 360
|
|
Db 372 CTGAGAACATCGTGAGGCGCAAAACGATGAGGTGCGGACCTATGATTTTGTGACAC 431
|
|
QY 361 TCAAGGTTACAGAGACCAAGGTGCTACCTCGCGGACGTTCTGTTGAGGACATC 420
|
|
Db 432 TCAAGGTTACAGAGACCAAGGTGCTACCTCGCGGACGTTCTGTTGAGGACATC 491
|
|
QY 421 TTGGTGTCTACAGCCAGAGATCCGGGACATGTTCCACCTGGGCCCTTCTGTGGACACC 480
|
|
Db 492 TTGGTGTCTACAGCCAGAGATCCGGGACATGTTCCACCTGGGCCCTTCTGTGGACACC 551
|
|
QY 481 GACTCCGACGTGAGCTGTCTCGAAGAGTTCTCCGGACGTGGCCGAGGAGGACCTG 540
|
|
Db 552 GACTCCGACGTGAGCTGTCTCGAAGAGTTCTCCGGACGTGGCCGAGGAGGACCTG 611
|
|
QY 541 GAGCAGATTCTGACGACGATACACCACTTGTGTGAAGCCGCTTCGAGAGATTTGCTGCTG 600
|
|
Db 612 GAGCAGATTCTGACGACGATACACCACTTGTGTGAAGCCGCTTCGAGAGATTTGCTGCTG 671
|
|
QY 601 CCGACAAAGAGATGCGGATGTGATCATCCACGAGAGATGACAAATATGTTGGCATC 660
|
|
Db 672 CCGACAAAGAGATGCGGATGTGATCATCCACGAGAGATGACAAATATGTTGGCATC 731
|
|
QY 661 AACCTGATCGTGACGACATCCAG 684
|
|
Db 732 AACCTGATCGTGACGACATCCAG 755
|
|
RESULT 4 725 bp mRNA linear EST 17-JUL-2001
B1256928
LOCUS 602975534F1 NIH_MGC_12 Homo sapiens CDNA clone IMAGE:5114816 5',
DEFINITION mRNA sequence.
B1256928
VERSION B1256928.1 GI:14811813
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 725)
NIH-MGC http://mgi.nci.nih.gov/.
TITLES National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LINL at:
<http://image.lnl.gov>
 Plate: LHAM11279 row: c column: 09
 High quality sequence stop: 707.

FEATURES
 source
 1..725
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5114816"
 /tissue_type="cervical carcinoma cell line"
 /lab_host="DH10B"
 /clone_lib="NIH MGC 12"
 /note="Organ: cervix; Vector: pCMV-Sport6; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally. Primer: oligo dt.
 Average insert size 1.4 kb. Library prepared by Life
 Technologies."

BASE COUNT 175 a 181 c 224 g 145 t
 ORIGIN

Query Match 81.2%; Score 677; DB 12; Length 725;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 677; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 104 GGAAGTGCACCGCTGTGAGAGATCATGAGTGTGGGACAGAGAGGTGGAAGC 163
 DB 1 GGAAGTGCACCGCTGTGAGAGATCATGAGTGTGGGACAGAGAGGTGGAAGC 60
 QY 164 GGCAGCGGAAGGTGATCTCTGAGCAGAGCAGGTTCTACAGAGTCTTGACGAGC 223
 DB 61 GGCAGCGGAAGGTGATCTCTGAGCAGAGCAGGTTCTACAGAGTCTTGACGAGC 120
 QY 224 AGAAGCCCAAGGCTTGAAGGACAGTCAATTTTACCATCCAGATGCTTTGATATG 283
 DB 121 AGAAGCCCAAGGCTTGAAGGACAGTCAATTTTACCATCCAGATGCTTTGATATG 180
 QY 284 ATTGATGACACGAGCTCTGAAGAACATCGTGAAGGCAAAAGGTGAGGTGCGACT 343
 DB 181 ATTGATGACACGAGCTCTGAAGAACATCGTGAAGGCAAAAGGTGAGGTGCGACT 240
 QY 344 ATGATTTTGTGACACTCAAGGTTTACAGAGACACGAGTGTCTACCTCGGACGTGG 403
 DB 241 ATGATTTTGTGACACTCAAGGTTTACAGAGACACGAGTGTCTACCTCGGACGTGG 300
 QY 404 TTCTGTTTGAAGGATCTTGTTCTTACAGCCAGAGATCCGGACATGTTCCACTGC 463
 DB 301 TTCTGTTTGAAGGATCTTGTTCTTACAGCCAGAGATCCGGACATGTTCCACTGC 360
 QY 464 GCTCTTCGTGAGACCGACTCCGAGCTCAGGCTGTCTCGAAGAGTTCTCCGGGACGTGC 523
 DB 361 GCTCTTCGTGAGACCGACTCCGAGCTCAGGCTGTCTCGAAGAGTTCTCCGGGACGTGC 420
 QY 524 GCCAGGGAAGGACCTGAGCAGATTCTGACGAGTACACCACTTGTGTAAGCCGAGCT 583
 DB 421 GCCAGGGAAGGACCTGAGCAGATTCTGACGAGTACACCACTTGTGTAAGCCGAGCT 480
 QY 584 TCGAGAGTTTGTGCTGCTGCGCAAAAGAGTATGCCGATGTATATCCACAGAGATGG 643
 DB 481 TCGAGAGTTTGTGCTGCTGCGCAAAAGAGTATGCCGATGTATATCCACAGAGATGG 540
 QY 644 ACAATATGTTGATCATCACTGATCGTGAAGCAGCATCCAGAGACATTTGTAATGTGACA 703
 DB 541 ACAATATGTTGATCATCACTGATCGTGAAGCAGCATCCAGAGACATTTGTAATGTGACA 600
 QY 704 TCTGCAATGACACGAGAGAGGTCCATGAGGCGGAGCTACAGCGGACCTTTTCTAGC 763
 DB 601 TCTGCAATGACACGAGAGAGGTCCATGAGGCGGAGCTACAGCGGACCTTTTCTAGC 660
 QY 764 CAGGGGACACCTCTGGG 780
 DB 661 CAGGGGACACCTCTGGG 677

RESULT 5
 BG826894
 LOCUS

DEFINITION BG826894 916 bp mRNA linear EST 22-MAY-2001
 602750978F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:4903819 5',
 mRNA sequence.

ACCESSION BG826894
 VERSION BG826894
 KEYWORDS BG826894.1 GI:14174481
 SOURCE EST.
 ORGANISM Homo sapiens (human)

REFERENCE Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 916)

AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>
 NATIONAL INSTITUTES OF HEALTH, Mammalian Gene Collection (MGC)
 UNPUBLISHED
 CONTACT: Robert Strausberg, Ph.D.
 Email: cgabs@mail.nih.gov
 Tissue Procurement: ATCC

COMMENT CDNA Library Preparation: Ling Hong/Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LINL at:
<http://image.lnl.gov>
 Plate: LHCMI802 row: k column: 20
 High quality sequence stop: 843.

FEATURES

source

1..916
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4903819"
 /tissue_type="rhodomyosarcoma"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC 17"
 /note="Organ: muscle; Vector: pOTB7; Site 1: EcoRI;
 Site 2: XhoI; cDNA made by oligo-dT priming.
 Directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGCACGAG(G). Size-selected >500bp
 for average insert size 1.8kb. Library constructed by
 Ling Hong in the laboratory of Gerald M. Rubin (University
 of California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 203 a 239 c 309 g 165 t
 ORIGIN

Query Match 77.9%; Score 650; DB 12; Length 916;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 650; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCTTGGCGGAGCGGAGACCTGCGAGAGCCCCGCGCGGAGCCGACCGTCCGAC 60
 DB 80 ATGGCTTGGCGGAGCGGAGACCTGCGAGAGCCCCGCGCGGAGCCGACCGTCCGAC 139
 QY 61 CAGCGGCCCTTCTGTAAGGGGTGAAGCGGCGGACCTCGAGGGGAAGTGAACCGGTCT 120
 DB 140 CAGCGGCCCTTCTGTAAGGGGTGAAGCGGCGGACCTCGAGGGGAAGTGAACCGGTCT 199
 QY 121 GAGAAATCATGAGATTGCTGGGACGAAGAGAGTGAACAGCGGCGAGCGGAAGTGTCT 180
 DB 200 GAGAAATCATGAGATTGCTGGGACGAAGAGAGTGAACAGCGGCGAGCGGAAGTGTCT 259
 QY 181 ATCTGAGCCAGGACAGGTTTCAAGAGTCTGACGCGCAGAGCAGAGGCAAGGCTTTG 240
 DB 260 ATCTGAGCCAGGACAGGTTTCAAGAGTCTGACGCGCAGAGCAGAGGCAAGGCTTTG 319
 QY 241 AAGGACAGTCAATTTTGAACATCCAGATGCTTTGATATGATTTGATGACAGAGCT 300
 DB 320 AAGGACAGTCAATTTTGAACATCCAGATGCTTTGATATGATTTGATGACAGAGCT 379
 QY 301 CTGAAGAAATCGTGAAGGCAAAAGGTGAGGTCCGACCTATGATTTTGTGACACAC 360

Db 380 CTGAGAACATCGTGGAGGCAAAACGCTGAGAGTGCCGACCTATGATTTGTGACACAC 439
 Qy 361 TCAAGTTACCAAGAGACCAAGGTGTCTTACCTCGGACGTGTCTTGTGAGGGATC 420
 Db 440 TCAAGGTTACCAAGAGACCAAGGTGTCTTACCTCGGACGTGTCTTGTGAGGGATC 499
 Qy 421 TTGGTGTTCACAGCCAGAGATCCGGACATGTTCCAGCTGGCCCTTCCTGAGACAC 480
 Db 500 TTGGTGTTCACAGCCAGAGATCCGGACATGTTCCAGCTGGCCCTTCCTGAGACAC 559
 Qy 481 GACTCCGACGTCAGGCTGTCTCGAAGATTCTCCGAGACGTGCGCCGAGGAGGACCTG 540
 Db 560 GACTCCGACGTCAGGCTGTCTCGAAGATTCTCCGAGACGTGCGCCGAGGAGGACCTG 619
 Qy 541 GAGCAGATTCTGACGACGATACCACTTGTGTGAAGCCGCTTCGAGAGATTCTGCTG 600
 Db 620 GACGACGTTCTGACGACGATACCACTTGTGTGAAGCCGCTTCGAGAGATTCTGCTG 679
 Qy 601 CCGACAAAGAGATGCGATGTGATCATCCACAGAGATGACAAATAT 650
 Db 680 CCGACAAAGAGATGCGATGTGATCATCCACAGAGATGACAAATAT 729
 RESULT 6
 LOCUS BX394295 1201 bp mRNA linear EST 13-MAY-2003
 DEFINITION BX394295 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
 ACCESSION BX394295
 VERSION BX394295.1 GI:30624219
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 1201)
 L1, W.B., Gruber, C., Jesse, J. and Polayes, D.
 Pull-length cDNA libraries and normalization
 Unpublished
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 4968.r For
 more information about this cluster, see
 http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CS0DC013AC07QPL1cluster=4968.r. Contact :
 Peng Liang Email : filiang@lifetech.com/ Invitrogen Corporation 1600
 Faraday Avenue Genoscope sequence ID : CS0DC013AC07QPL1.
 Location/Qualifiers
 1..1201
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DC013IE13"
 /tissue_type="NEUROBLASTOMA COT 25-NORMALIZED"
 /clone_id="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-clio(dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and cloned into the Not I and EcoR V
 sites of the pCMVSPORT 6 vector. Library was normalized."
 BASE COUNT 288 a 284 c 363 g 226 t 40 others
 ORIGIN
 Query Match 77.1% Score 643; DB 13; Length 1201;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 693; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 46 GCCGACCGTCCGACACGACGCGCCCTTCTGATAGGGGTGAGCGGCGCACTGCGACGGG 105
 Db 132 GCCGACCGTCCGACACGACGCGCCCTTCTGATAGGGGTGAGCGGCGCACTGCGACGGG 131

Qy 106 AAGTCGACCGTGTGAGAGAGATCATGAGTTGCTGGGACAGAACGAGGTGAAACGCG 165
 Db 192 AAGTCGACCGTGTGAGAGAGATCATGAGTTGCTGGGACAGAACGAGGTGAAACGCG 251
 Qy 166 CAGCGAAGGTGTGATCTCTGAGCCAGACAGAGTTCTTACAAAGTCTCTGACGAGACAG 225
 Db 252 CAGCGAAGGTGTGATCTCTGAGCCAGACAGAGTTCTTACAAAGTCTCTGAGGAGACAG 311
 Qy 226 AAGCCGACGCTTGAAGAGACATGAAATTTTGAACATCCAGATCCCTTGAATGAT 285
 Db 312 AAGCCGACGCTTGAAGAGACATGAAATTTTGAACATCCAGATCCCTTGAATGAT 371
 Qy 286 TTGATGACACAGACTCTTGAAGAAATCTGTGAGAGGAGGAGGAGGAGGAGGAGGAGG 345
 Db 372 TTGATGACACAGACTCTTGAAGAAATCTGTGAGAGGAGGAGGAGGAGGAGGAGGAGG 431
 Qy 346 GATTTGTGACACACTCAAGTTACAGAGACCAAGGAGGAGGAGGAGGAGGAGGAGGAGG 405
 Db 432 GATTTGTGACACACTCAAGTTACAGAGACCAAGGAGGAGGAGGAGGAGGAGGAGGAGG 491
 Qy 406 CTGTTGAGGAGCATCTTGTGTCTACAGCCAGAGATCCGAGACATGTTCCACTGCGC 465
 Db 492 CTGTTGAGGAGCATCTTGTGTCTACAGCCAGAGATCCGAGACATGTTCCACTGCGC 551
 Qy 466 CTCTTGTGTGACACCACTCCGACCTCCAGCTGCTCTCGAAGAGTTCTCCGGGACGTGCG 525
 Db 552 CTCTTGTGTGACACCACTCCGACCTCCAGCTGCTCTCGAAGAGTTCTCCGGGACGTGCG 611
 Qy 526 CGAGGAGGAGCATCTGAGAGATTTGACGACGATACCACTCTCTGAGAGCGGCGCTTC 585
 Db 612 CGAGGAGGAGCATCTGAGAGATTTGACGACGATACCACTCTCTGAGAGCGGCGCTTC 671
 Qy 586 GAGGAGTTCTGCTGCTGCGCAAAAGATATGCCGATGTATATCCACGAGAGGTGAC 645
 Db 672 GAGGAGTTCTGCTGCTGCGCAAAAGATATGCCGATGTATATCCACGAGAGGTGAC 731
 Qy 646 AATATGTTGCCATCAACTGATCTGTCGACGACATCCAGACATTTGATGATGATC 705
 Db 732 AATATGTTGCCATCAACTGATCTGTCGACGACATCCAGACATTTGATGATGATC 791
 Qy 706 TGCAATGACACCGAGAGGATCCAAATGGGCGGA 739
 Db 792 TGCAATGACACCGAGAGGATCCAAATGGGCGGA 825
 RESULT 7
 LOCUS BG770518 730 bp mRNA linear EST 15-MAY-2001
 DEFINITION BG770518 602734226F1 NIH_MGC_49 Homo sapiens cDNA clone IMAGE:4859577 5',
 mRNA sequence.
 ACCESSION BG770518 GI:14081171
 VERSION BG770518
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 730)
 NIH-MGC http://img.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished
 Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-r@mail.nih.gov
 Tissue Procurement: ATCC/DCTD/DTP
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (JLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/JLNL at:
 http://image.jnl.gov
 Plate: LLCM1715 row: h column: 10
 High quality sequence stop: 720.

FEATURES
source

Location/Qualifiers
1. .730
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4859577"
/tissue_type="melanotic melanoma, high MDR (cell line)"
/lab_host="DH10B (phage-resistant)"
/clone_1ib="NIH MGC 49"
/note="Organ: skin; Vector: pORF7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dt priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

BASE COUNT 156 a 190 c 258 g 126 t

Query Match 76.7%; Score 640; DB 12; Length 730;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 640; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCTTCGGGGGAGGCGAAGACTGCGAGAGCCCGCGCGGAGCCGACCGTCCGAC 60
DB 82 ATGGCTTCGGGGGAGGCGAAGACTGCGAGAGCCCGCGCGGAGCCGACCGTCCGAC 141
QY 61 CAGCGGCCCCCTCCGATAGGGGTGAGCGCGGCACTGCGACCGGAAAGTCCAGCTGTGT 120
DB 142 CAGCGGCCCCCTCCGATAGGGGTGAGCGCGGCACTGCGACCGGAAAGTCCAGCTGTGT 201
QY 121 GAGAAGATCATGAGTTCCTGGGACAGAACGAGGTGAAACGCGGACGCGGAGGTGCTC 180
DB 202 GAGAAGATCATGAGTTCCTGGGACAGAACGAGGTGAAACGCGGACGCGGAGGTGCTC 261
QY 181 ATCTGAGCCAGGACAGGTTCTCAAGGTCTCTGACGCGAGACAGAAAGCCAGGCTTG 240
DB 262 ATCTGAGCCAGGACAGGTTCTCAAGGTCTCTGACGCGAGACAGAAAGCCAGGCTTG 321
QY 241 AAAAGACAGTCAATTTTGAACATCCAGATCCCTTTGATTAATGATTTGATGACAGACT 300
DB 322 AAAAGACAGTCAATTTTGAACATCCAGATCCCTTTGATTAATGATTTGATGACAGACT 381
QY 301 CTGAAGAATCATGAGTTCGAGAGGCGAAACGCTGAGGTGCGGACCTATGATTTTGACAC 360
DB 382 CTGAAGAATCATGAGTTCGAGAGGCGAAACGCTGAGGTGCGGACCTATGATTTTGACAC 441
QY 361 TCAAGGTTACCGAGACCCAGGTGCTCAACCTGCGGACGCTGTTCTGTTGAGGAGATC 420
DB 442 TCAAGGTTACCGAGACCCAGGTGCTCAACCTGCGGACGCTGTTCTGTTGAGGAGATC 501
QY 421 TTGGTGTTCACACGCGAGAGATCCGGGACATGTTCCACCTGCGGCTCTTGTGTGACACC 480
DB 502 TTGGTGTTCACACGCGAGAGATCCGGGACATGTTCCACCTGCGGCTCTTGTGTGACACC 561
QY 481 GACTCCAGCTCAGGCTGTCTCAAGAGTTCTCCGGGACGTGCGCGAGGAGGAGGACTG 540
DB 562 GACTCCAGCTCAGGCTGTCTCAAGAGTTCTCCGGGACGTGCGCGAGGAGGAGGACTG 621
QY 541 GAGCAGATTCTGACGACGATACCACTTCGTGAAGCCGCGCTTCAAGAGTTCTGCTG 600
DB 622 GAGCAGATTCTGACGACGATACCACTTCGTGAAGCCGCGCTTCAAGAGTTCTGCTG 681
QY 601 CCGACAAAGAGTATGCCGATGATCATCCCAAGAGAG 640
DB 682 CCGACAAAGAGTATGCCGATGATCATCCCAAGAGAG 721

RESULT 8
BQ719741
LOCUS BQ719741 956 bp mRNA linear EST 16-JUL-2002

DEFINITION
AGENCOURT 8229663 lupski dorsal root ganglion Homo sapiens cDNA
clone IMAGE:6184625 5', mRNA sequence.
BQ719741
VERSION BQ719741.1 GI:21858638
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 956)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: gsr@bbs.fda.nih.gov
Tissue Procurement: Dr. James R. Lupski
CDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
Plate: LLM3574 row: b column: 18
High quality sequence stop: 552.

FEATURES
source

Location/Qualifiers
1. .956
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6184625"
/sex="male"
/tissue_type="dorsal root ganglia"
/dev stage="adult, 36 yr"
/lab host="DH10B"
/clone_1ib="lupski dorsal root ganglion"
/note="Vector: pCMV-SPORT (Life Technologies); Site 1: NotI; Site 2: SalI; cDNA made by oligo-dt priming. Directionally cloned using the following adaptors: 5'-TCGACCCAGCGCTCG-3' and 5'-GACTAGTTCAGATCGGAGCGGCGGCTT(15)-3'. Size selected > 1 kb for average insert length 1.7 kb. This is a primary library, non-amplified. Library constructed by Life Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor College of Medicine) and is available through Life Technologies."

BASE COUNT 202 a 272 c 296 g 186 t

Query Match 74.2%; Score 619; DB 13; Length 956;
Best Local Similarity 100.0%; Pred. No. 2,3e-309;
Matches 619; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 GAGGCGAAGCTCGAGAGCCCGCGCGGAGGCGGACCGTCCGACCGAGCGGCGCTTCC 73
DB 20 GAGGCGAAGCTCGAGAGCCCGCGCGGAGGCGGACCGTCCGACCGAGCGGCGCTTCC 79
QY 74 TGATAGGGGTGAGCGGCGGACCTGCGAGCGGAGTGCACCGGTGTGAGAAATCATGG 133
DB 80 TGATAGGGGTGAGCGGCGGACCTGCGAGCGGAGTGCACCGGTGTGAGAAATCATGG 139
QY 134 AGTTGCTGGACAGAACGAGGTGAAACAGCGGACGCGAAGGTGTCATCTGAGCCAG 193
DB 140 AGTTGCTGGACAGAACGAGGTGAAACAGCGGACGCGAAGGTGTCATCTGAGCCAG 199
QY 194 ACAGGTTTCACAGGTCCTGACGCGGACGACAGAAAGCCCAAGGCTTGAAGAAGACATACA 253
DB 200 ACAGGTTTCACAGGTCCTGACGCGGACGACAGAAAGCCCAAGGCTTGAAGAAGACATACA 259
QY 254 ATTTGACCATCAGATGCTTTGATTAATGATTTGACAGAGACTTGAAGAATCATCG 313
DB 260 ATTTGACCATCAGATGCTTTGATTAATGATTTGACAGAGACTTGAAGAATCATCG 319
QY 314 TGAGGCGCAAAACGCTGAGGTGCGGACCTATGATTTGTGACACACTCAAGGTTACAG 373

Db 320 TGGAGGGCAAAACGCTGAGAGTGCCGACCTATGATTTTGTGACACACTCAAGGTTACG 379
Qy 374 AGACCAAGGTGCTTACCTCCGACCTGCTTCTGTTGAGGATCTTGTGTTCTACA 433
Db 380 AGACCAAGGTGCTTACCTCCGACCTGCTTCTGTTGAGGATCTTGTGTTCTACA 439
Qy 434 GCACAGAGATCCGGGACATGTTCCACCTGCGCCTCTTGAGACACGACCTCCGACGTC 493
Db 440 GCCAGAGATCCGGGACATGTTCCACCTGCGCCTCTTGAGACACGACCTCCGACGTC 499
Qy 494 GCGTCTCTGAAAGATTCTCCGAGCGTGCGCCGAGGAGAGGACCTTGAGACAGATTCTGA 553
Db 500 GCGTCTCTGAAAGATTCTCCGAGCGTGCGCCGAGGAGAGGACCTTGAGACAGATTCTGA 559
Qy 554 CGCAGTACACCACTTGTGTGAAGCCGCTTCGAGAGATTCTGCTCTCCGACAAAGAGT 613
Db 560 CGCAGTACACCACTTGTGTGAAGCCGCTTCGAGAGATTCTGCTCTCCGACAAAGAGT 619
Qy 614 ATGCCGATGATCATCC 632
Db 620 ATGCCGATGATCATCC 638

RESULT 9
BG390519 989 bp mRNA linear EST 12-MAR-2001
LOCUS 60241618.F1 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:4524308 5',
DEFINITION mRNA sequence.
ACCESSION BG390519
VERSION BG390519.1 GI:13283967
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 989)
NIH-MGC http://mgs.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LHAM0428 row: f column: 21
High quality sequence stop: 750.
Location/Qualifiers

FEATURES
source 1..989
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4524308"
/tissue_type="embryonal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_id="NIH_MGC_92"
/note="Organ: testis; Vector: pCMV-SPORT6; Site 1: NciI;
Site 2: SalI. Cloned unidirectionally; oligo-dT primed.
Average insert size 2.5 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

BASE COUNT 256 a 258 c 314 g 161 t
ORIGIN

Query Match 73.7%; Score 615; DB 10; Length 989;
Best Local Similarity 100.0%; Freq. No. 2.8e-307;
Matches 615; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 CGAAGACTGAGAGCCCGCGCGGACGACCTCCGACACAGGCGCCCTTCGTGAT 77

Db 1 CGAAGACTGAGAGCCCGCGCGGACGACCTCCGACACAGGCGCCCTTCGTGAT 60
Qy 78 AGGGGTGAGCGGGGACATCGCCAGCGGAAAGTGAACCTGTGTGAAAGATCTGAGATT 137
Db 61 AGGGGTGAGCGGGGACATCGCCAGCGGAAAGTGAACCTGTGTGAAAGATCTGAGATT 120
Qy 138 GCTGGAGCAAGACGAGGTGAAACAGCGGACGCGAAAGTGTATCTCTGAGCCAGACAG 197
Db 121 GCTGGAGCAAGACGAGGTGAAACAGCGGACGCGAAAGTGTATCTCTGAGCCAGACAG 180
Qy 198 GTTCTCAAGGTCCTGACGCGACAGAGCAAGGCTTGAAGAGACATTAATTT 257
Db 181 GTTCTCAAGGTCCTGACGCGACAGAGCAAGGCTTGAAGAGACATTAATTT 240
Qy 258 TGACCATCCAGATGCTTTGATATATATTTGATGACAGAGATCTGAAACATCTGTGA 317
Db 241 TGACCATCCAGATGCTTTGATATATATTTGATGACAGAGATCTGAAACATCTGTGA 300
Qy 318 GGGCAAAACGCTGAGAGTGCCGACCTATGATTTTGTGACACACTCAAGGTTACAGAGAC 377
Db 301 GGGCAAAACGCTGAGAGTGCCGACCTATGATTTTGTGACACACTCAAGGTTACAGAGAC 360
Qy 378 CACGCTGCTTACCTGCGGACGCTGCTTCTGTTGAGGACATCTGATCTTACAGCCA 437
Db 361 CACGCTGCTTACCTGCGGACGCTGCTTCTGTTGAGGACATCTGATCTTACAGCCA 420
Qy 438 GGAATCCGAGACATGTTCCACTGCGCTCTTCTGAGACACGACATCCGAGTCAAGGCT 497
Db 421 GGAATCCGAGACATGTTCCACTGCGCTCTTCTGAGACACGACATCCGAGTCAAGGCT 480
Qy 498 GTCTCAAGAGTCTCCGGGACGTGCGGAGGAGGACCTTGAGACAGATTCTGACGCA 557
Db 481 GTCTCAAGAGTCTCCGGGACGTGCGGAGGAGGACCTTGAGACAGATTCTGACGCA 540
Qy 558 GTACACCACTTCTGTAAGCGGCTTTCGAGAGATTCTGCTGCGCCGACAAAGAGATATGC 617
Db 541 GTACACCACTTCTGTAAGCGGCTTTCGAGAGATTCTGCTGCGCCGACAAAGAGATATGC 600
Qy 618 CGATGTGATCATCC 632
Db 601 CGATGTGATCATCC 615

RESULT 10
BG491384 869 bp mRNA linear EST 27-MAR-2001
LOCUS 602535670.F1 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:4684936 5',
DEFINITION mRNA sequence.
ACCESSION BG491384
VERSION BG491384.1 GI:13452896
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 869)
NIH-MGC http://mgs.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: DCTD/DTF
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LHCMI493 row: k column: 17
High quality sequence stop: 863.
Location/Qualifiers

FEATURES
source 1..869

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_image="6684936"
/tissue_type="amelanotic melanoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_1ib="N1H MGC 41"
/note="Organ: skin; Vector: pCRB7; Site 1: XhoI; Site 2:
EcoRI; CDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
N1H MGC library."
BASE COUNT      197 a      223 c      291 g      156 t      2 others
ORIGIN

Query Match      73.3%; Score 611; DB 10; Length 869;
Best Local Similarity 100.0%; Pred. No. 3.2e-305;
Matches 611; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Query Match	73.3%	Score 611	DB 10	Length 869
Best Local Similarity	100.0%	Pred. No. 3.2e-305		
Matches	611	Conservative	0	Mismatches 0; Indels 0; Gaps 0;
Qy	1	ATGGCTTGCGCGGAGCGAGACTGCCAGAGCCCGCGCGCGAGGCCCGACCGTCCGAC	60	
Db	52	ATGGCTTCGGCGGGAGCGGAGACTGCCAGAGCCCGCGCGCGAGGCCCGACCGTCCGAC	111	
Qy	61	CAGCGGCCCTTCTGATATGAGGTGAGCGCGCGCACTGCCAGCGGGAATGACCGTGTGT	120	
Db	112	CAGCGGCCCTTCTGATATGAGGTGAGCGCGCGCACTGCCAGCGGGAATGACCGTGTGT	171	
Qy	121	GAGAAAGATCATGAGAGTCTGCGGACAGAAAGAGGTGAAACAGCGGCAGCGGAAGTGTCTC	180	
Db	172	GAGAAAGATCATGAGAGTCTGCGGACAGAAACAGAGGTGAAACAGCGGCAGCGGAAGTGTCTC	231	
Qy	181	ATCTGAGCCAGGACAGAGTCTCAAGGTCTCTGACGCGAGAGCAAGAGCCGCTTG	240	
Db	232	ATCTGAGCCAGGACAGAGTCTCAAGGTCTCTGACGCGAGAGCAAGAGCCGCTTG	291	
Qy	241	AAAGGACGTCATATTTGACCATCCAGATGCGTTTGAATAATTTGATGCACAGACT	300	
Db	292	AAAGGACGTCATATTTGACCATCCAGATGCGTTTGAATAATTTGATGCACAGACT	351	
Qy	301	CTGAAGAACATCGTGGAGGGCAAAACGCTGAGAGTGCAGACCTATGATTTTGTACACAC	360	
Db	352	CTGAAGAACATCGTGGAGGGCAAAACGCTGAGAGTGCAGACCTATGATTTTGTACACAC	411	
Qy	361	TCAAAGTTTCCAGAGACCAACGCTGTCTACCTCTGCGACGTTTCTGTTTGAAGGCAATC	420	
Db	412	TCAAAGTTTCCAGAGACCAACGCTGTCTACCTCTGCGACGTTTCTGTTTGAAGGCAATC	471	
Qy	421	TTGGTGTTCTACAGCAGAGACATCCGGGACATGTTTCCACTGCGGCTTCTGTGTGACACC	480	
Db	472	TTGGTGTTCTACAGCAGAGACATCCGGGACATGTTTCCACTGCGGCTTCTGTGTGACACC	531	
Qy	481	GACTCCGACGTCAAGCTGTCTTCGAAAGATTCTCCGGGACGTCGCGCCGACGAGAGGACCTG	540	
Db	532	GACTCCGACGTCAAGCTGTCTTCGAAAGATTCTCCGGGACGTCGCGCCGACGAGAGGACCTG	591	
Qy	541	GAGCAGATTTGAGCGCAGTACACCACTTCTGTGAAGCCGCGCTTGTGAGAGATTCTGCTGT	600	
Db	592	GAGCAGATTTGAGCGCAGTACACCACTTCTGTGAAGCCGCGCTTGTGAGAGATTCTGCTGT	651	
Qy	601	CCGACCAAGAA 611		
Db	652	CCGACCAAGAA 662		

RESULT 11

AUI31406 779 bp mRNA linear EST 01-AUG-2002

LOCUS AUI31406 NT2RP3 Homo sapiens cDNA clone NT2RP3002519 5', mRNA

DEFINITION

ACCESSION AUI31406

VERSION	AU11406.1	GI:10991760
KEYWORDS	EST.	
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.	
AUTHORS	Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saio,K., Kawai,Y., Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and Isogai,T.	
TITLE	HRI human cDNA project	
JOURNAL	Unpublished	
COMMENT	Contact: Takao Isogai Genomics Laboratory Helix Research Institute 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan Tel: 81-438-52-3975 Fax: 81-438-52-3986 Email: genomics@hri.co.jp HRI human cDNA project; 5'- & 3'-end one pass sequencing; Helix Research Institute; cDNA library construction; Department of Virology, Institute of Medical Science, University of Tokyo, and Helix Research Institute.	
FEATURES	Location/Qualifiers	
source	1..779	
	/organism="Homo sapiens"	
	/mol_type="mRNA"	
	/db_xref="taxon:9606"	
	/clone="NT2RP3002519"	
	/cell_type="teratocarcinoma"	
	/cell_line="NT2"	
	/clone_id="NT2RP3"	
	/note="Vector: pME18SFL3; mRNA from NT2 neuronal precursor cells after 2-weeks retinoic acid (RA) induction"	
BASE COUNT	179 a 197 c 253 g 147 t 3 others	
ORIGIN		
Query Match	72.5%; Score 605; DB 9; Length 779;	
Best Local Similarity	99.8%; Pred. No. 3.9e-302;	
Matches	655; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	
1	ATGGCTTGGGGGAGGCGGAGACTGCGCAGAGCCCGCGCGGAGCCGACCGCTCCGCAC	60
24	ATGGCTTGGGGGAGGCGGAGACTGCGCAGAGCCCGCGCGGAGCCGACCGCTCCGCAC	83
61	CAGCGGCCCTTCTCTGATATAGGGGTAGCGCGGCACTGCGACGGGAAATCGACCGTGT	120
84	CAGCGGCCCTTCTCTGATATAGGGGTAGCGCGGCACTGCGACGGGAAATCGACCGTGT	143
121	GAGAAAGATCATGAGATTCTGCGGACAGAACGAGGTGGAACAGCGGACGGAAAGGTGTC	180
144	GAGAAAGATCATGAGATTCTGCGGACAGAACGAGGTGGAACAGCGGACGGAAAGGTGTC	203
181	ATCTGAGCCGAGACAGGTTCTTACAAGGTCTTGACGCGGACGAGAGAGGCCAAGGCTTG	240
204	ATCTGAGCCGAGACAGGTTCTTACAAGGTCTTGACGCGGACGAGAGAGGCCAAGGCTTG	263
241	AAAGGACGTCATATTTTGAACATCCAGATGCGCTTGTATATGATTTGATGACACAGACT	300
264	AAAGGACGTCATATTTTGAACATCCAGATGCGCTTGTATATGATTTGATGACACAGACT	323
301	CTGAAGAACATCTGTGAGGGCAAAACGCTGAGGTGCGACCTTATGATTTTGTGACACAC	360
324	CTGAAGAACATCTGTGAGGGCAAAACGCTGAGGTGCGACCTTATGATTTTGTGACACAC	383
361	TCAGAGTTACAGAGACCAACGCTGATCTACCTCGGACGCTGTTCTGTTGAGGGCATC	420
384	TCAGAGTTACAGAGACCAACGCTGATCTACCTCGGACGCTGTTCTGTTGAGGGCATC	443
421	TTGGTGTCTACAGCAGGAGATCCGGGACATGTTCACCTCGGCGCTTGTGTGACACAC	480
444	TTGGTGTCTACAGCAGGAGATCCGGGACATGTTCACCTCGGCGCTTGTGTGACACAC	503

QY 481 GACTCCGACGTGAGTGTCTCCAGAGATTCTCCGGACCTGCGCGAGGAGGACCTG 540
DB 504 GACTCCGACGTGAGTGTCTCCAGAGATTCTCCGGACCTGCGCGAGGAGGACCTG 553
QY 541 GAGCAGATTCTGACGAGTACACCACTTCTGTAAGCCGGCTTCCAGAGATTCTGCTG 600
DB 564 GAGCAGATTCTGACGAGTACACCACTTCTGTAAGCCGGCTTCCAGAGATTCTGCTG 623
QY 601 CCGACAAAGAGATGCGGATGATCATCCGACGAGAGTGAACATATGTTGC 656
DB 624 CCGACAAAGAGATGCGGATGATCATCCGACGAGAGTGAACATATGTTGC 679

RESULT 12
LOCUS B0072501 1036 bp mRNA linear EST 02-APR-2002
DEFINITION AGENCOURT_6838909 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:5761694
5', mRNA sequence.
ACCESSION B0072501
VERSION B0072501.1 GI:19901547
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1036)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: LLM12810 row: 1 column: 15
High quality sequence stop: 644.

FEATURES
source
1. 1036
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5761694"
/lab_host="DH10B"
/clone_1lb="NIH_MGC_122"
/note="Organ: pooled lung and spleen; Vector: pCMV-SPORT6; site_1: NotI; Site 2: EcoRV (destroyed); RNA source anonymous pool of 24 week female lung, 16 week female spleen, and 20-22 week male spleens. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.4 kb, insert size range 1-3 kb. Library is normalised and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 026. Note: this is a NIH_MGC Library."

BASE COUNT 221 a 285 c 346 g 184 t
ORIGIN
Query Match 72.2%; Score 602; DB 13; Length 1036;
Best Local Similarity 100.0%; Pred. No. 1.6e-300; Indels 0; Gaps 0;
Matches 602; Conservative 0; Mismatches 0;

QY 1 ATGGCTTCGCGGAGCGGAGAGCTGCGAGAGCCCGCGCGAGGCGGACCGTCCGAC 60
DB 64 ATGGCTTCGCGGAGCGGAGAGCTGCGAGAGCCCGCGCGAGGCGGACCGTCCGAC 123
QY 61 CAGCGGCGCTTCTGATAGGGGTGAGCGGCGGCACTGCGGAGAAAGTGCACCGTGTG 120
DB 124 CAGCGGCGCTTCTGATAGGGGTGAGCGGCGGCACTGCGGAGAAAGTGCACCGTGTG 183

QY 121 GAGAGATCATGAGTGTGCTGGGACAGAAAGAGTGAACAGCGGACGCGAAGTGTGTC 180
DB 184 GAGAGATCATGAGTGTGCTGGGACAGAAAGAGTGAACAGCGGACGCGAAGTGTGTC 243
QY 181 ATCTGAGCCGAGACAGGTTCTTCAAGGTCCTGACGCGACAGAGCAAGGCCAAGCCCTTG 240
DB 244 ATCTGAGCCGAGACAGGTTCTTCAAGGTCCTGACGCGACAGAGCAAGGCCAAGCCCTTG 303
QY 241 AAAGGACATTAATTTTGAATCCATCCAGATGCTTGAATTAATTAATTTGATGCAAGACT 300
DB 304 AAAGGACATTAATTTTGAATCCATCCAGATGCTTGAATTAATTAATTTGATGCAAGACT 363
QY 301 CTGAGAGACATGCTGGAGGGCAAAACGCTGAGAGTCCGACCTATGATTTTGGACACAC 360
DB 364 CTGAGAGACATGCTGGAGGGCAAAACGCTGAGAGTCCGACCTATGATTTTGGACACAC 423
QY 361 TCAAGGTTACAGAGACACAGGTGCTTACCTGCGAGAGTGTGTTGTTGAGGCAATC 420
DB 424 TCAAGGTTACAGAGACACAGGTGCTTACCTGCGAGAGTGTGTTGTTGAGGCAATC 483
QY 421 TTGGTCTTCAACAGCCAGAGATCCGGGACATGTTCCACTGCGCTCTTCTGTGACACC 480
DB 484 TTGGTCTTCAACAGCCAGAGATCCGGGACATGTTCCACTGCGCTCTTCTGTGACACC 543
QY 481 GACTCCGACGTGAGTGTCTCGAAGAGTTCTCCGGAGCTGCGCGAGGAGGACCTG 540
DB 544 GACTCCGACGTGAGTGTCTCGAAGAGTTCTCCGGAGCTGCGCGAGGAGGACCTG 603
QY 541 GAGCAGATTCTGACGAGTACACCACTTCTGTAAGCCGGCTTCCAGAGATTCTGCTG 600
DB 604 GAGCAGATTCTGACGAGTACACCACTTCTGTAAGCCGGCTTCCAGAGATTCTGCTG 663
QY 601 CC 602
DB 664 CC 665

RESULT 13
LOCUS B1258532 848 bp mRNA linear EST 17-JUL-2001
DEFINITION 602972340F1 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:5112032 5',
mRNA sequence.
ACCESSION B1258532
VERSION B1258532.1 GI:14814971
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 848)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: LLM11271 row: 0 column: 09
High quality sequence stop: 839.

FEATURES
source
1. 848
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5112032"
/issue_type="cervical carcinoma cell line"
/lab_host="DH10B"
/clone_1lb="NIH_MGC_12"

/note="Organ: cervix; Vector: PCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.4 kb. Library prepared by Life
Technologies."

BASE COUNT

196 a 219 c 271 g 162 t

ORIGIN

Query Match 70.6%; Score 589; DB 12; Length 848;

Best Local Similarity 100.0%; Pred. No. 7.9e-294; Mismatches 0; Indels 0; Gaps 0;

Matches 589; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

108 GTGACCGGTGTGTGAGAGATCATGAGTTGCTGGGACAGAACGAGGTGAAACAGCGGCA 167
86 GTGACCGGTGTGTGAGAGATCATGAGTTGCTGGGACAGAACGAGGTGAAACAGCGGCA 145
168 GCGGAAGTGTGTCTCTGAGCCGAGACAGTTCTTACAGAGTCTTACAGCGGACGCGAA 227
146 GCGGAAGTGTGTCTCTGAGCCGAGACAGTTCTTACAGAGTCTTACAGCGGACGCGAA 205
228 GCGGAAGTGTGTCTCTGAGCCGAGACAGTTCTTACAGAGTCTTACAGCGGACGCGAA 287
206 GCGGAAGTGTGTCTCTGAGCCGAGACAGTTCTTACAGAGTCTTACAGCGGACGCGAA 265
288 GATGACAGGACTCTGAGAGACATCTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 347
266 GATGACAGGACTCTGAGAGACATCTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 325
348 TTTTGTGACACATCAAGGTTTACAGAGACAGCGGTGTCTTACCGTGGAGCGTGTCT 407
326 TTTTGTGACACATCAAGGTTTACAGAGACAGCGGTGTCTTACCGTGGAGCGTGTCT 385
408 GTTGAAGGAGACTCTGTGTCTTACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 467
386 GTTGAAGGAGACTCTGTGTCTTACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 445
468 CTTCGTGAGACACGAGCTCCGACGTCAAGGCTGTCTGAGAGTCTCCGAGCGGAGCGG 527
446 CTTCGTGAGACACGAGCTCCGACGTCAAGGCTGTCTGAGAGTCTCCGAGCGGAGCGG 505
528 AGGAGAGGAGCTTGTGAGACAGATTTGACGCGAGTACACCACTTCTGAGAGCGGCTTGA 587
506 AGGAGAGGAGCTTGTGAGACAGATTTGACGCGAGTACACCACTTCTGAGAGCGGCTTGA 565
588 GAGATTCTGCTGCTGCGACAAAGAGTATGCGGATGTATCATCCAGAGAGAGTGA 647
566 GAGATTCTGCTGCTGCGACAAAGAGTATGCGGATGTATCATCCAGAGAGAGTGA 625
648 TATGTTGACATCAACCTGATCTGAGACATTCAGAGACATTTGTAAT 696
626 TATGTTGACATCAACCTGATCTGAGACATTCAGAGACATTTGTAAT 674

RESULT 14

BM545603 1055 bp mRNA linear EST 20-FEB-2002

LOCUS AGENCOURT 6500279 NIH_MGC_125 Homo sapiens cDNA clone IMAGE:5587864

DEFINITION 5', mRNA sequence.

ACCESSION BM545603.1 GI:18777841

VERSION EST.

KEYWORDS Homo sapiens (human)

SOURCE Homo sapiens

ORGANISM Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE 1 (bases 1 to 1055)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)

COMMENT Unpublished

Contact: Robert Strauberg, Ph.D.

Email: cgapds-f@mail.nih.gov

DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LINL at:
<http://image.lnl.gov>
Plate: L14M12358 row: a column: 17
High quality sequence stop: 579.
Location/Qualifiers
1..1055

FEATURES

source

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:5587864"

/lab_host="DH10B"

/clone_lib="NIH MGC 125"

/note="Organ: ovary (pool of 3); Vector: PCMV-SPORT6;

Site 1: EcoRV (destroyed); Site 2: NotI; RNA source pool

of three ovaries, from females ranging in age from 38 to

49 yo. Library is oligo-dr primed and directionally cloned

(EcoRV site is destroyed upon cloning). Average insert

size 2.1 kb, insert size range 1-3.5 kb. Library is

normalized and enriched for full-length clones and was

constructed by C. Gruber (Invitrogen). Research Genetics

tracking code 036."

BASE COUNT

208 a 343 c 314 g 190 t

ORIGIN

Query Match 69.8%; Score 582; DB 12; Length 1055;

Best Local Similarity 100.0%; Pred. No. 3.6e-290; Mismatches 0; Indels 0; Gaps 0;

Matches 582; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGCGTTCGCGGAGGAGCGAAGACTGAGAGCGCGCGCGCGGAGCGCGCGCGCGCGCGCG 60
49 ATGCGTTCGCGGAGGAGCGAAGACTGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 108
61 CAGCGCGCGCTTCTCTGATGAGGAGTGAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120
109 CAGCGCGCGCTTCTCTGATGAGGAGTGAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 168
121 GAGAAGTATGAGAGTTGCTGGGACAGAACGAGGTGGAACAGCGGCGGCGGCGGCGGCGGCGG 180
169 GAGAAGTATGAGAGTTGCTGGGACAGAACGAGGTGGAACAGCGGCGGCGGCGGCGGCGGCGG 228
181 ATCTGAGCGAGGAGCGGTTCTTCAAGGTCCTTACGCGGCGGAGCGGAGCGGCGGCGGCGG 240
229 ATCTGAGCGAGGAGCGGTTCTTCAAGGTCCTTACGCGGCGGAGCGGAGCGGCGGCGGCGG 288
241 AAAGGACAGTACATTTTGAACATCCAGATGCTTTGATTAATGATTTGATGACAGAGCT 300
289 AAAGGACAGTACATTTTGAACATCCAGATGCTTTGATTAATGATTTGATGACAGAGCT 348
301 CTGAGAGACATCGTGAAGGCGAAAGCGTGAAGTCCGACCTATATGATTTTGTGACACAC 360
349 CTGAGAGACATCGTGAAGGCGAAAGCGTGAAGTCCGACCTATATGATTTTGTGACACAC 408
361 TCAAGTTACAGAGACACAGGTTCTTCAAGGTCCTTACGCGGAGCGGAGTGTGTTGTTGAGGCGATC 420
409 TCAAGTTACAGAGACACAGGTTCTTCAAGGTCCTTACGCGGAGCGGAGTGTGTTGTTGAGGCGATC 468
421 TTGTTGTTTACAGCGAGGAGATCCGGAATGCTTCACTGAGGCTCTTCTGTGACAC 480
469 TTGTTGTTTACAGCGAGGAGATCCGGAATGCTTCACTGAGGCTCTTCTGTGACAC 528
481 GATTCGAGGTCAGGCTGTCTTGAAGAGTTTCCGGAAGTGTGCGCGAGGAGGAGGAGGAGGAGGAG 540
529 GATTCGAGGTCAGGCTGTCTTGAAGAGTTTCCGGAAGTGTGCGCGAGGAGGAGGAGGAGGAGGAG 588
541 GAGCAGATTCTGAGCGAGTACACCACTTCTGTAAGCGGCGGCGGCGGCGGCGGCGGCGGCGG 582
589 GAGCAGATTCTGAGCGAGTACACCACTTCTGTAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 630

RESULT 15

BI830878

This Page Blank (uspto)


```
DB 141 LeuValhethrYsergIngluilearYAspMethetiSleuArgLeuPheValAspThr 160
QY 481 GACTCCGACGTCAGCGTGTCTCGAGA 507
DB 161 AsperAspValArgLeuSerArgArg 169

RESULT 2
Q9QY68 PRELIMINARY; PRT; 105 AA.
ID Q9QY68
AC Q9QY68
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Uridine Kinase (Fragment).
GN UK.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SD; TISSUE=Spinal cord;
RX MEDLINE=20050059; PubMed=10581173;
RA Yuh I., Yaoi T., Matsunabe S., Okajima S., Hirasawa Y., Fushiki S.;
RL Biochem. Biophys. Res. Commun. 266:104-109(1999).
DR EMBL; AB030700; BAA83085.1; -.
DR InterPro; IPR006083; PRK URK.
DR InterPro; IPR000764; Uridine_kin.
DR Pfam; PF00465; PRK.1.
DR PRINTS; PR00988; URIDINKINASE.
KW Kinase.
FT NON_TER 1 105
SQ SEQUENCE 105 AA; 11967 MW; 9DB43C15E07EB029 CRC64;

Alignment Scores:
Pred. No.: 1.23e-16 Length: 105
Score: 25.00 Matches: 25
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 9.03% Indels: 0
DB: 11 Gaps: 0

US-09-896-522-3 (1-834) x Q9QY68 (1-105)
QY 565 ACCTTCGTGAAGCGGCTTCGAGAGATTCTGCCTGCCGACAAAGAAGTATGCCGATGTG 624
DB 36 ThrPheValIysProAlaPheGluGluPheCysLeuProThrIlystyAlaAspVal 55
QY 625 ATCATCCACGAGAGA 639
DB 56 IleIleProArgIly 60

RESULT 3
Q9BU42 PRELIMINARY; PRT; 111 AA.
ID Q9BU42
AC Q9BU42;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Similar to uridine monophosphate kinase.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RA Strauberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
```

```
DR EMBL; BC002906; AAH02906.1; -.
DR InterPro; IPR006083; PRK URK.
DR InterPro; IPR000764; Uridine_kin.
DR Pfam; PF00465; PRK.1.
DR PRINTS; PR00988; URIDINKINASE.
KW Kinase.
SQ SEQUENCE 111 AA; 12587 MW; E6688B1B86F432A9 CRC64;

Alignment Scores:
Pred. No.: 1.23e-16 Length: 111
Score: 25.00 Matches: 25
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 9.03% Indels: 0
DB: 4 Gaps: 0

US-09-896-522-3 (1-834) x Q9BU42 (1-111)
QY 565 ACCTTCGTGAAGCGGCTTCGAGAGATTCTGCCTGCCGACAAAGAAGTATGCCGATGTG 624
DB 37 ThrPheValIysProAlaPheGluGluPheCysLeuProThrIlystyAlaAspVal 56
QY 625 ATCATCCACGAGAGA 639
DB 57 IleIleProArgIly 61

RESULT 4
Q8C476 PRELIMINARY; PRT; 261 AA.
ID Q8C476
AC Q8C476;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Uridine-cytidine Kinase 2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium.
RA The RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RL 60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
DR EMBL; AK082837; BAC38646.1; -.
SQ SEQUENCE 261 AA; 29438 MW; 780AA3C298AA8153 CRC64;

Alignment Scores:
Pred. No.: 1.14e-16 Length: 261
Score: 25.00 Matches: 25
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 9.03% Indels: 0
DB: 11 Gaps: 0

US-09-896-522-3 (1-834) x Q8C476 (1-261)
QY 565 ACCTTCGTGAAGCGGCTTCGAGAGATTCTGCCTGCCGACAAAGAAGTATGCCGATGTG 624
DB 187 ThrPheValIysProAlaPheGluGluPheCysLeuProThrIlystyAlaAspVal 206
QY 625 ATCATCCACGAGAGA 639
DB 207 IleIleProArgIly 211

RESULT 5
Q92528 PRELIMINARY; PRT; 111 AA.
ID Q92528
AC Q92528;
DT 01-FEB-1997 (TREMBlrel. 02, Created)
```

DT 01-FBB-1997 (Tremblrel. 02, last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, last annotation update)
DE 5'-terminal region of UMK.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=96411689; PubMed=8812458;
RA Oaki K., Kuroki T., Hayashi S., Nakamura Y.;
RT "Isolation of three testis-specific genes (TSA303, TSA806, TSA903) by
RT a differential mRNA display method.";
RL Genomics 36:316-319 (1996).
DR EMBL; D78335; BA11349.1; -;
DR InterPro; IPR006083; PRK_URK.
DR Pfam; PF00485; PRK_1.
DR PRINTS; PR00988; URIDINKINASE.
SQ SEQUENCE 111 AA; 12617 MW; 81F6E4019CF428AB CRC64;

Alignment Scores:
Pred. No.: 3.41e-06 Length: 111
Score: 15.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.42% Indels: 0
DB: 4 Gaps: 0

US-09-896-522-3 (1-834) x Q92528 (1-111)

QY 565 ACCTGCGAGCGCGCTCGAGAGTTCGCTGCGGACAAAG 609
DB 37 TnpheVallysProAlaPheGluPheCysLeuProThrlys 51

RESULT 6

OSMRJ1 PRELIMINARY; PRT; 260 AA.
AC OSMRJ1;
DT 01-OCT-2002 (Tremblrel. 22, Created)
DT 01-OCT-2002 (Tremblrel. 22, last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, last annotation update)
DE ID13909p.
GN CG6364.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkely;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Fiske E.,
RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouanevong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celinker S.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
RT EMBL; AY119583; AA50237.1; -;
DR FLYBase; FBgn0039179; CG6364.
DR InterPro; IPR006083; PRK_URK.
DR InterPro; IPR000764; Uridine_kin.
DR Pfam; PF00485; PRK_1.
DR PRINTS; PR00988; URIDINKINASE.
SQ SEQUENCE 260 AA; 29223 MW; 03CAA6DCA04A16E5 CRC64;

Alignment Scores:
Pred. No.: 3.18e-06 Length: 260
Score: 15.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0

Query Match: 5.42% Indels: 0
DB: 5 Gaps: 0

US-09-896-522-3 (1-834) x OSMRJ1 (1-260)

QY 388 TACCTGCGAGCGTGTGTTTGAGGCGATCTGTGTTCTAC 432
DB 135 TyrProAlaAspValValLeuPheGluGlyIleLeuValPheTyr 149

Search completed: November 25, 2003, 07:58:09
Job time: 40.5285 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 25, 2003, 00:22:12 ; Search time 3240.32 Seconds

(without alignments)
10529.410 Million cell updates/sec

Title: US-09-896-522-3

Sequence: 1 atggtctcgcggagcgca.....ccagcagcagaccacctga 834

Scoring table: OLIGO_NUC

Searched: 2888711 seqs, 20454813386 residues

word size : 15

Total number of hits satisfying chosen parameters: 29369

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Maximum DB seq length: 200000000000

Post-processing: Listing first 45 summaries

Database :

1: gb_ba:*
2: gb_hng:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pac:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_ser:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_ov:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_hng_hum:*
31: em_hng_inv:*
32: em_hng_other:*
33: em_hng_mus:*
34: em_hng_pln:*
35: em_hng_rad:*
36: em_hng_mam:*
37: em_hng_vrt:*
38: em_sy:*
39: em_hngo_hum:*
40: em_hngo_mus:*
41: em_hngo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	834	100.0	834	6	AX449219	AX449219 Sequence
2	834	100.0	1624	6	AX449217	AX449217 Sequence
3	834	100.0	2160	6	BD157613	BD157613 Primer fo
4	834	100.0	2160	9	AC022317	AC022317 Homo sapi
5	783	93.9	834	9	AF254133	AF254133 Homo sapi
6	783	93.9	1022	9	AF237290	AF237290 Homo sapi
7	575	68.9	755	6	BD146824	BD146824 Primer fo
8	517	62.0	1395	9	AF125106	AF125106 Homo sapi
9	508	60.9	2072	9	BC015547	BC015547 Homo sapi
10	497	59.6	2228	9	AK057848	AK057848 Homo sapi
11	338	40.5	734	6	AX540411	AX540411 Sequence
12	183	21.9	147492	9	AL358781	AL358781 Human DNA
13	36	4.3	660	11	BV077668	BV077668 S212P6227
14	36	4.3	1810	10	MUSURK1	L31783 Mus musculu
15	36	4.3	1959	10	BC025146	BC025146 Mus muscu
16	36	4.3	233210	9	AC097693	AC097693 Rattus no
17	36	4.3	241882	10	AL808027	AL808027 Mouse DNA
18	36	4.3	244105	2	AC098897	AC098897 Rattus no
19	36	4.3	254677	2	AC068494	AC068494 Mus muscu
20	33	4.0	254993	9	AC078885	AC078885 Mus muscu
21	23	2.8	336	9	BT006860	BT006860 Homo sapi
22	23	2.8	336	12	BT007558	BT007548 Synthetic
23	23	2.8	383	11	G41891	G41891 SHGC-35183
24	23	2.8	744	9	AB062451	AB062451 Homo sapi
25	23	2.8	1129	9	AF236637	AF236637 Homo sapi
26	23	2.8	1209	9	BC002906	BC002906 Homo sapi
27	23	2.8	1322	6	AX135546	AX135546 Sequence
28	23	2.8	99866	9	AL358115	AL358115 Human DNA
29	23	2.8	164246	2	AC016371	AC016371 Homo sapi
30	22	2.6	233137	2	AC127758	AC127758 Rattus no
31	21	2.5	841	11	BV017303	BV017303 S212P6359
32	21	2.5	952	10	MUSURK1A	L31784 Mus musculu
33	21	2.5	3060	4	OC010358	U10358 OryctoIagus
34	21	2.5	3060	6	I13759	I13759 Sequence 1
35	21	2.5	192336	2	AC118474	AC118474 Mus muscu
36	21	2.5	211075	10	AC076974	AC076974 Mus muscu
37	20	2.4	655	11	BV077340	BV077340 S212P6009
38	20	2.4	1581	5	AF004318	AF004318 Carassius
39	20	2.4	1784	5	BC045968	BC045968 Danio rer
40	20	2.4	4663	1	AY181035	AY181035 Methyloba
41	20	2.4	62229	2	AC100224	AC100224 Mus muscu
42	20	2.4	68200	2	AC009801_3	Continuation (4 of
43	20	2.4	133036	2	AC101821	AC101821 Mus muscu
44	20	2.4	148865	2	AL954171	AL954171 Danio rer
45	20	2.4	162978	2	AC132083	AC132083 Mus muscu

ALIGNMENTS

RESULT 1	AX449219	LOCUS	AX449219	834 bp	DNA	linear	PAT 03-JUL-2002
DEFINITION	Sequence 3 from Patent WO0202761.	ACCESSION	AX449219				
VERSION	AX449219.1	GI:	21697996				
KEYWORDS							
SOURCE							
ORGANISM	Homo sapiens (human)						
	Homo sapiens						
	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;						
	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.						
REFERENCE	1						
AUTHORS	Glucksmann, M.A.						
TITLE	57658, a human uridine kinase and uses thereof						
JOURNAL	Patent: WO 0202761-A 3 10-JAN-2002;						

Millennium Pharmaceuticals, Inc. (US)
Location/Qualifiers
1. .834
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 194 a 218 c 265 g 157 t
ORIGIN

Query Match 100.0%; Score 834; DB 6; Length 834;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 834; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCTTCGGGGGAGGAGAGACTCGGAGAGCCCCCGGCGGAGGCGGACCGTCGGCAC 60
DB 1 ATGGCTTCGGGGGAGGAGAGACTCGGAGAGCCCCCGGCGGAGGCGGACCGTCGGCAC 60
QY 61 CAGCGGCCCTTCCTGATAGGGGTAGCGCGGCACTGCGCAGCGGGAAGTCGACCGTGTGT 120
DB 61 CAGCGGCCCTTCCTGATAGGGGTAGCGCGGCACTGCGCAGCGGGAAGTCGACCGTGTGT 120
QY 121 GAGAGATCATGAGATTGCTGGGACAGAACGAGGTGAAACGCGGACGCGAAGGTGTCT 180
DB 121 GAGAGATCATGAGATTGCTGGGACAGAACGAGGTGAAACGCGGACGCGAAGGTGTCT 180
QY 181 ATCTGAGCCGAGACAGGTTCTACAAGTCTCTGACGAGAGCAGAAAGGCCAAGCCCTTG 240
DB 181 ATCTGAGCCGAGACAGGTTCTACAAGTCTCTGACGAGAGCAGAAAGGCCAAGCCCTTG 240
QY 241 AAAGGACAGTCAATTTTGAACATCCAGATCCCTTTGATATGATTGATGACAGAGACT 300
DB 241 AAAGGACAGTCAATTTTGAACATCCAGATCCCTTTGATATGATTGATGACAGAGACT 300
QY 301 CTGAAGAACATCGTGGAGGGCAAAACGCTGAGAGTCCGACCTTATGATTTTGTGACAC 360
DB 301 CTGAAGAACATCGTGGAGGGCAAAACGCTGAGAGTCCGACCTTATGATTTTGTGACAC 360
QY 361 TCAAGGTTACGAGAGACACGCGTGTCTACCCCTCGGACGCTGTTGTTGAGGGATC 420
DB 361 TCAAGGTTACGAGAGACACGCGTGTCTACCCCTCGGACGCTGTTGTTGAGGGATC 420
QY 421 TTGGTGTCTACAGCCAGAGATCCGGGACATGTTTCCACTGCGCCTCTTCTGTGACACC 480
DB 421 TTGGTGTCTACAGCCAGAGATCCGGGACATGTTTCCACTGCGCCTCTTCTGTGACACC 480
QY 481 GACTCCGACGTCAGGCTGTCTCGAAGAGTTCTCCGGGACGTGCGCCGAGGGAGGACCTG 540
DB 481 GACTCCGACGTCAGGCTGTCTCGAAGAGTTCTCCGGGACGTGCGCCGAGGGAGGACCTG 540
QY 541 GAGGAGATTTCGAGGAGGACATACCAACCTTCGTGAAAGCGGCGCTTCGAGAGTTCTG 600
DB 541 GAGGAGATTTCGAGGAGGACATACCAACCTTCGTGAAAGCGGCGCTTCGAGAGTTCTG 600
QY 601 CCGACAAAGAGTATGCGATGTATCATCCACGAGAGTGAACAATATGTTGCCATC 660
DB 601 CCGACAAAGAGTATGCGATGTATCATCCACGAGAGTGAACAATATGTTGCCATC 660
QY 661 AACCTGATCGTGCAGACATCCAGACATTTGAAATGTTGACATCTGCAATGTCACCGA 720
DB 661 AACCTGATCGTGCAGACATCCAGACATTTGAAATGTTGACATCTGCAATGTCACCGA 720
QY 721 GGAAGGTCATGAGGCGGAGCTACAAGGCGACCTTTCTGAGCAGAGGGGACACCTGGG 780
DB 721 GGAAGGTCATGAGGCGGAGCTACAAGGCGACCTTTCTGAGCAGAGGGGACACCTGGG 780
QY 781 ATGCTGACCTCTGGCAACGCTCACATTTGAGATCTCAGACAGACCCCACTGA 834
DB 781 ATGCTGACCTCTGGCAACGCTCACATTTGAGATCTCAGACAGACCCCACTGA 834

RESULT 2
AX449217
LOCUS AX449217 1624 bp DNA linear PAT 03-JUN-2002

DEFINITION Sequence 1 from Patent WO0202761.
ACCESSION AX449217
VERSION AX449217.1 GI:21697994
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Glucksmann, M.A.
TITLE 57658, a human uridine kinase and uses thereof
JOURNAL Patent: WO 0202761-A 1 10-JAN-2002;
Millennium Pharmaceuticals, Inc. (US)
FEATURES
Source
Location/Qualifiers
1. .1624
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
94..927
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD37771.1"
/db_xref="GI:21697995"
/translation="MASAGBEDCESPAPADRPHORPLIGVSGTASGKSTVCEKIM
ELIGONEVBOBORXVILISODREFKVLTAEOAKALKGOYRNDPDAFPNDLMHRLTK
NIVEGTVFVPTDPTVHSRLPETTVVPAADVLPFGILVFSOEIRDMFHLRLPVDY
DSVRLSRVLRDVRGRDLEOILTYOTTFVAPAREFCLPKRYADVILIPRGVNDV
ALNLIYOHIDILINDICMKHREGSNGRSYKRTFSEPDHPMLISGKSHLESSRP
H"
BASE COUNT 354 a 427 c 506 g 337 t
ORIGIN

Query Match 100.0%; Score 834; DB 6; Length 1624;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 834; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCTTCGGGGGAGGAGAGACTCGGAGAGCCCCCGGAGGCGGACCGTCGGCAC 60
DB 94 ATGGCTTCGGGGGAGGAGAGACTCGGAGAGCCCCCGGAGGCGGACCGTCGGCAC 153
QY 61 CAGCGGCCCTTCCTGATAGGGGTAGCGCGGCACTGCGCAGCGGGAAGTCGACCGTGTGT 120
DB 61 CAGCGGCCCTTCCTGATAGGGGTAGCGCGGCACTGCGCAGCGGGAAGTCGACCGTGTGT 120
QY 121 GAGAGATCATGAGATTGCTGGGACAGAACGAGGTGAAACGCGGACGCGAAGGTGTCT 180
DB 121 GAGAGATCATGAGATTGCTGGGACAGAACGAGGTGAAACGCGGACGCGAAGGTGTCT 180
QY 181 ATCTGAGCCGAGACAGGTTCTACAAGTCTCTGACGAGAGCAGAAAGGCCAAGCCCTTG 240
DB 181 ATCTGAGCCGAGACAGGTTCTACAAGTCTCTGACGAGAGCAGAAAGGCCAAGCCCTTG 240
QY 241 AAAGGACAGTCAATTTTGAACATCCAGATCCCTTTGATATGATTGATGACAGAGACT 300
DB 241 AAAGGACAGTCAATTTTGAACATCCAGATCCCTTTGATATGATTGATGACAGAGACT 300
QY 301 CTGAAGAACATCGTGGAGGGCAAAACGCTGAGAGTCCGACCTTATGATTTTGTGACAC 360
DB 301 CTGAAGAACATCGTGGAGGGCAAAACGCTGAGAGTCCGACCTTATGATTTTGTGACAC 360
QY 361 TCAAGGTTACGAGAGACACGCGTGTCTACCCCTCGGACGCGTGTCTGTTTGAAGGCATC 420
DB 361 TCAAGGTTACGAGAGACACGCGTGTCTACCCCTCGGACGCGTGTCTGTTTGAAGGCATC 420
QY 421 TTGGTGTCTACAGCCAGAGATCCGGGACATGTTTCCACTGCGCCTCTTCTGTGACACC 480
DB 421 TTGGTGTCTACAGCCAGAGATCCGGGACATGTTTCCACTGCGCCTCTTCTGTGACACC 480
QY 481 GACTCCGACGTCAGGCTGTCTCGAAGAGTTCTCCGGGAGTGTGCGGAGGAGGACCTG 540
DB 481 GACTCCGACGTCAGGCTGTCTCGAAGAGTTCTCCGGGAGTGTGCGGAGGAGGACCTG 540
QY 541 GAGGAGATTTCGAGGAGGACATACCAACCTTCGTGAAAGCGGCGCTTCGAGAGTTCTG 600
DB 541 GAGGAGATTTCGAGGAGGACATACCAACCTTCGTGAAAGCGGCGCTTCGAGAGTTCTG 600

Db 634 GAGCAATTCGACGAGTACACACCTTGTAAGCGCGCTTCGAGAGTCTGCTG 693
Qy 601 CCGACAAAGAAAGATGCCGATGATCATGCCACGAGAGTGAACAATATGTTGCCATC 660
Db 694 CCGACAAAGAAAGATGCCGATGATCATGCCACGAGAGTGAACAATATGTTGCCATC 753
Qy 661 AACCTGATCGTGAGCAGCATCCAGACATCTGAATGTTGACATCTGCAATGAGCAGCA 720
Db 754 AACCTGATCGTGAGCAGCATCCAGACATCTGAATGTTGACATCTGCAATGAGCAGCA 813
Qy 721 GAGGGTCCAAATGCGGAGCTCAAGCGGACCTTTTCTGAGCCAGGAGCAACCTGGG 780
Db 814 GAGGGTCCAAATGCGGAGCTCAAGCGGACCTTTTCTGAGCCAGGAGCAACCTGGG 873
Qy 781 ATGCTGACCTTGCGCAACCGTCAACATTTGGAGTCCAGAGCAGACCCCACTGA 834
Db 874 ATGCTGACCTTGCGCAACCGTCAACATTTGGAGTCCAGAGCAGACCCCACTGA 927
RESULT 3
BD157613 2160 bp DNA linear PAT 17-JAN-2003
LOCUS BD157613
DEFINITION Primer for synthesizing full-length cDNA and use thereof.
ACCESSION BD157613
VERSION BD157613.1 GI:27863371
KEYWORDS JP 2002191363-A/12456.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 2160)
AUTHORS Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
Ishii,S., Sugiyama,T., Makamatsu,A., Nagai,K. and Otsuki,T.
TITLE Primer for synthesizing full-length cDNA and use thereof
JOURNAL HELIX RESEARCH INSTITUTE
COMMENT OS Homo sapiens (human)
PN JP 2002191363-A/12456
PD 09-JUL-2002 JP 2000280990
FP 28-JUL-2000 JP 2000280990
PI TOSHIO OTA,TAKAO ISOGAI,TETSUO NISHIKAWA,KOJI HAYASHI,KOORU
PI SAITO,
PI JUNICHI YAMAMOTO,SHIZUKO ISHII,TOMOYASU SUGIYAMA,AI MAKAMATSU,
PI KEIICHI NAGAI,TETSUJI OTSUKI
PC C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/ PC
10,
PC C12P21/02,C12Q1/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof FH Key
FEATURES
source Location/Qualifiers
FT CDS
1.2160
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 457 a 591 c 671 g 441 t
ORIGIN
Query Match 100.0%; Score 834; DB 6; Length 2160;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 834; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 215 GAGAAGATCATGAGATTGCTGGGACAGAAAGGTGAACAAGCGGACGGAAAGGTGTC 274
Qy 181 ATCTTGAAGCAGAGACAGGTTTCTAACAAGTCTTGAACGCGAGACAGAAAGCCAGGCTTG 240
Db 275 ATCTTGAAGCAGAGACAGGTTTCTAACAAGTCTTGAACGCGAGACAGAAAGCCAGGCTTG 334
Qy 241 AAAGCAGTACAAATTTTGAACCATCCAGATGCTTGTATATGATTGATGACAGAGACT 300
Db 335 AAAGCAGTACAAATTTTGAACCATCCAGATGCTTGTATATGATTGATGACAGAGACT 394
Qy 301 CTGAAGAACATCGTGGAGGGCAAAAAGGTGAGGTGCGAAGCTATGATTTTGTGACACAC 360
Db 395 CTGAAGAACATCGTGGAGGGCAAAAAGGTGAGGTGCGAAGCTATGATTTTGTGACACAC 454
Qy 361 TCAGAGTTACAGAGACACAGTGTCTACCTCGGAGAGTGTCTTTTGAAGGCAATC 420
Db 455 TCAGAGTTACAGAGACACAGTGTCTACCTCGGAGAGTGTCTTTTGAAGGCAATC 514
Qy 421 TTGGTGTCTACAGCGAGAGATCCGGGACATGTTCCACTGGGCTTTCGTGACACC 480
Db 515 TTGGTGTCTACAGCGAGAGATCCGGGACATGTTCCACTGGGCTTTCGTGACACC 574
Qy 481 GACTCCGAGCTGAGGCTGTCTCGAAGAGTTCTCCGAGAGTGTGCGGAGGAGGAGCTG 540
Db 575 GACTCCGAGCTGAGGCTGTCTCGAAGAGTTCTCCGAGAGTGTGCGGAGGAGGAGCTG 634
Qy 541 GAGCAGATTCTGACCGAGTACACACCTTCGTGAGACCGGCTTTCGAGAGATTCTGCTG 600
Db 635 GAGCAGATTCTGACCGAGTACACACCTTCGTGAGACCGGCTTTCGAGAGATTCTGCTG 694
Qy 601 CCGACAAAGAAAGATGCCGATGATCATCCACGAGAGTGAACAATATGTTGCCATC 660
Db 695 CCGACAAAGAAAGATGCCGATGATCATCCACGAGAGTGAACAATATGTTGCCATC 754
Qy 661 AACCTGATCGTGAGCAGCATCCAGACATCTGAATGTTGACATCTGCAATGAGCAGCA 720
Db 755 AACCTGATCGTGAGCAGCATCCAGACATCTGAATGTTGACATCTGCAATGAGCAGCA 814
Qy 721 GAGGGTCCAAATGCGGAGCTCAAGCGGACCTTTTCTGAGCCAGGAGCAACCTGGG 780
Db 815 GAGGGTCCAAATGCGGAGCTCAAGCGGACCTTTTCTGAGCCAGGAGCAACCTGGG 874
Qy 781 ATGCTGACCTTGCGCAACCGTCAACATTTGGAGTCCAGAGCAGACCCCACTGA 834
Db 875 ATGCTGACCTTGCGCAACCGTCAACATTTGGAGTCCAGAGCAGACCCCACTGA 928
RESULT 4
AK022317 2160 bp mRNA linear PRI 01-AUG-2002
LOCUS AK022317
DEFINITION Homo sapiens cDNA FLJ12255 f1s, clone MAMMAL1001476, highly similar
to URIDINE KINASE (EC 2.7.1.48).
ACCESSION AK022317
VERSION AK022317.1 GI:10433687
KEYWORDS oligo capping, f1s (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
AUTHORS Isogai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y.,
Nishikawa,T., Nagai,K., Sugano,S., Shiratori,A., Sudo,H.,
Wagatsuma,M., Hosoi,T., Kaku,Y., Kodaira,H., Kondo,H.,
Sugawara,M., Takahashi,M., Chiba,Y., Ishida,S., Murakawa,K.,
Ono,Y., Takiguchi,S., Watanabe,S., Kimura,K., Murakami,K.,
Ishii,S., Kawai,Y., Saito,K., Yamamoto,J., Makamatsu,A.,
Makamura,Y., Nagahara,K., Masuo,Y., Ninomiya,K. and Iwayanagi,T.
TITLE NEDO human cDNA sequencing project
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2160)
AUTHORS Isogai,T. and Otsuki,T.
TITLE Direct Submission

JOURNAL Submitted (23-AUG-2000) Takao Isegai, Helix Research Institute, Genomics Laboratory/ 1532-3 Yana, Kisarazu, Chiba 252-0812, Japan (E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)

COMMENT NEDO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing; Research Association for Biotechnology; cDNA library construction, 5'- & 3'-end one pass sequencing and clone selection; Helix Research Institute (supported by Japan Key Technology Center etc.) and Department of Virology, Institute of Medical Science, University of Tokyo.

FEATURES
source
1. 2160
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MAMMA101476"
/tissue_type="Mammary gland"
/clone_id="MAMMA1"
/note="Cloning vector: pME18SFL3"
95..928
/note="unnamed protein product"
/codon_start=1
/protein_id="BAB14010.1"
/db_xref="GI:10433688"
/translation="MASAGDEDCSPAPADPHORPFLIGVSGTASGKSTVCEKIM ELLEGNEVEORORRVVILSDRFYKVLTAECOKAKALKGYNFDPDPAFNDLMDHRTLK NIVEGKTEVPYDFTVTHSRILPETTVVYADVLVFEGLIVFYSOEIRDMFHLFVDT DSDVRLSRVLDVRGRDLEQILQYTTTVKPAPEERCLPTKKYADVILPPGYDMVY AINLIVQHIQDILNIDICMKHMGSGNSRYSKRTFSEPDHGMILTSGRSHLESSSRP H"

BASE COUNT 457 a 591 c 671 g 441 t

ORIGIN

Query Match 100.0%; Score 834; DB 9; Length 2160; Best Local Similarity 100.0%; Pred. No. 0; Matches 834; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGGCTTCGGCGGAGCGAAGACTGGAGAGCCCGCGCGGACCGTCCGAC 60
95 ATGGCTTCGGCGGAGCGAAGACTGGAGAGCCCGCGCGGACCGTCCGAC 154
61 CAGCGGCCCTTCCTGATAGGGGTGAGCGCGGCACTGCCAGCGGAAGTGCACCGTGTGT 120
155 CAGCGGCCCTTCCTGATAGGGGTGAGCGCGGCACTGCCAGCGGAAGTGCACCGTGTGT 214
121 GAGAAATCATGAGTTCCTGGGACAGAACGAGGTGAAACGCGGACCGGAAGTGTGT 180
215 GAGAAATCATGAGTTCCTGGGACAGAACGAGGTGAAACGCGGACCGGAAGTGTGT 274
181 ATCTGAGCCAGGACAGTTCACAGGTCTGACGCGACAGACGAAAGCCAGCCCTTG 240
275 ATCTGAGCCAGGACAGTTCACAGGTCTGACGCGACAGACGAAAGCCAGCCCTTG 334
241 AAAGGACAGTACATTTTGAACCATCCAGATGCTTTGTAATGATTTGATGACAGGACT 300
335 AAAGGACAGTACATTTTGAACCATCCAGATGCTTTGTAATGATTTGATGACAGGACT 394
301 CTGAAAGACATGTCGAGAGGCAAAACGTCGAGGTGCGACCTATGATTTTGTGACACAC 360
395 CTGAAAGACATGTCGAGAGGCAAAACGTCGAGGTGCGACCTATGATTTTGTGACACAC 454
361 TCAGAGTTTACAGAGACCAAGGTGTCTACCTCTCGGACGAGGTGTCTTTGTTGAGGATC 420
455 TCAGAGTTTACAGAGACCAAGGTGTCTACCTCTCGGACGAGGTGTCTTTGTTGAGGATC 514
421 TTGGTGTCTTACACGACGAGAGATCCGGGACATGTTTCACTGCGCCCTCTTCTGTGACACC 480
515 TTGGTGTCTTACACGACGAGAGATCCGGGACATGTTTCACTGCGCCCTCTTCTGTGACACC 574
481 GACTCCGACGTCAGGCTGTCTCGAAGATTCTCCGGGACGTGCGCGAGGAGGAGGACCTG 540
575 GACTCCGACGTCAGGCTGTCTCGAAGATTCTCCGGGACGTGCGCGAGGAGGAGGACCTG 634

QY 541 GAGCAGATTCTGACGCGAGTACACCACTCTTGTAAGCCGCGCTTGGAGGAGTTCTGCTG 600
DB 635 GAGCAGATTCTGACGCGAGTACACCACTCTTGTAAGCCGCGCTTGGAGGAGTTCTGCTG 694
QY 601 CCGACAAAGAGTATCCGATGATCATCCACGAGAGTGGACAAATATGTTGCTATC 660
DB 695 CCGACAAAGAGTATCCGATGATCATCCACGAGAGTGGACAAATATGTTGCTATC 754
QY 661 AACCTGATGTCGACGACATCCAGGACATTTGTAATGTCGACATCTGCAATATGACCGA 720
DB 755 AACCTGATGTCGACGACATCCAGGACATTTGTAATGTCGACATCTGCAATATGACCGA 814
QY 721 GAGAGGTCGAATGAGGAGGAGTCAAGCGGACCTTTCTGAGCCAGGAGGACCGCTGGG 780
DB 815 GAGAGGTCGAATGAGGAGGAGTCAAGCGGACCTTTCTGAGCCAGGAGGACCGCTGGG 874
QY 781 ATGCTGACCTCTGGCAAGCGTGCACATTTGAGTCCAGACGAGACCCCACTGA 834
DB 875 ATGCTGACCTCTGGCAAGCGTGCACATTTGAGTCCAGACGAGACCCCACTGA 928

RESULT 5
AF254133 834 bp mRNA linear PRI 02-MAY-2001
LOCUS AF254133 Homo sapiens uridine kinase mRNA, complete cds.
DEFINITION AF254133.1 GI:13924749
ACCESSION AF254133.1 GI:13924749
VERSION
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 834)
Ho, Y.S. and Johnson, R.K.
Human uridine kinase from prostate cancer cell line (LNCap)
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 834)
Ho, Y.S. and Johnson, R.K.
Direct Submission
Submitted (10-APR-2000) Oncology Research, SmithKline Beecham, 709 Swedeland Road, King of Prussia, PA 19406, USA
JOURNAL
TITLE
AUTHORS
REFERENCES
FEATURES
source
1. 834
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1..834
/note="protein kinase"
/codon_start=1
/product="uridine kinase"
/protein_id="AK49122.1"
/db_xref="GI:13924750"
/translation="MASAGDEDCSPAPADPHORPFLIGVSGTASGKSTVCEKIM ELLEGNEVEORORRVVILSDRFYKVLTAECOKAKALKGYNFDPDPAFNDLMDHRTLK NIVEGKTEVPYDFTVTHSRILPETTVVYADVLVFEGLIVFYSOEIRDMFHLFVDT DSDVRLSRVLDVRGRDLEQILQYTTTVKPAPEERCLPTKKYADVILPPGYDMVY AINLIVQHIQDILNIDICMKHMGSGNSRYSKRTFSEPDHGMILTSGRSHLESSSRP H"

BASE COUNT 194 a 219 c 265 g 156 t

ORIGIN

Query Match 93.9%; Score 783; DB 9; Length 834; Best Local Similarity 99.9%; Pred. No. 0; Matches 833; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 ATGGCTTCGGCGGAGCGAAGACTGGAGAGCCCGCGCGGACCGACCGTCCGAC 60
DB 1 ATGGCTTCGGCGGAGCGAAGACTGGAGAGCCCGCGCGGACCGACCGTCCGAC 60
QY 61 CAGCGGCCCTTCCTGATAGGGGTGAGCGCGGCACTGCCAGCGGAAGTGCACCGTGTGT 120
DB 61 CAGCGGCCCTTCCTGATAGGGGTGAGCGCGGCACTGCCAGCGGAAGTGCACCGTGTGT 120

QY	121	GAGAAAGATCATGAGAACTTGTGGGACAGAAACGAGGTGGAAACAGCGGACAGCGAAAGGTGTGC	180
Db	121	GAGAAAGATCATGAGAACTTGTGGGACAGAAACGAGGTGGAAACAGCGGACAGCGAAAGGTGTGC	180
QY	181	ATCTCGAACCCAGSACAGGTTCTTACAAAGGTCCTGACGGGACAGCAAGAGCCCAAGGCTTTG	240
Db	181	ATCTCGAACCCAGSACAGGTTCTTACAAAGGTCCTGACGGGACAGCAAGAGCCCAAGGCTTTG	240
QY	241	AAAGGACAGTACATTTTGGATCCATCCGATGCGCTTTTGATATGATTTGATGACAGAGACT	300
Db	241	AAAGGACAGTACATTTTGGATCCATCCGATGCGCTTTTGATATGATTTGATGACAGAGACT	300
QY	301	CTGAAGAACATCTGTGGAGGGCAAAACGGTGGAGGTGGCCGACCTATGATTTTGTGACACAC	360
Db	301	CTGAAGAACATCTGTGGAGGGCAAAACGGTGGAGGTGGCCGACCTATGATTTTGTGACACAC	360
QY	361	TCAAAGTTTACCGAGAACACCGGTGGTCTTACCCCTGGGAGACGGTGTCTGTGTTTGAAGGACATC	420
Db	361	TCAAAGTTTACCGAGAACACCGGTGGTCTTACCCCTGGGAGACGGTGTCTGTGTTTGAAGGACATC	420
QY	421	TTGGTGTTCACAGCCAGAGATCCGGGACATGTTTCACTGCGGCTCTTTCGTGGACACC	480
Db	421	TTGGTGTTCACAGCCAGAGATCCGGGACATGTTTCACTGCGGCTCTTTCGTGGACACC	480
QY	481	GACTCCGACGTCAGGCTGTCTTCTGAAGAGATTCTCCGGGACGTGCGCCGAGGAGGGACCTG	540
Db	481	GACTCCGACGTCAGGCTGTCTTCTGAAGAGATTCTCCGGGACGTGCGCCGAGGAGGGACCTG	540
QY	541	GAGCAGATTCTGACGCGACAGTACACACCTTCGTAAGACCGGACCTTGAGAGAGTTCTTGCCCTG	600
Db	541	GAGCAGATTCTGACGCGACAGTACACACCTTCGTAAGACCGGACCTTGAGAGAGTTCTTGCCCTG	600
QY	601	CCGACAAAGAAAGTATCCGATGTGATCAATCCACGAGAGATGGAACATATGTGTTGCCATC	660
Db	601	CCGACAAAGAAAGTATCCGATGTGATCAATCCACGAGAGATGGAACATATGTGTTGCCATC	660
QY	661	AACTGATCTGTGACAGACATCCAGACATTTCTGAATGTTGACATCTGCAATGGCACCGA	720
Db	661	AACTGATCTGTGACAGACATCCAGACATTTCTGAATGTTGACATCTGCAATGGCACCGA	720
QY	721	GAGGGGTCCAAATGGGGCGGAGCTPACAAAGGGGACCTTTCTGAGCGAGGGGACCAACCTGGG	780
Db	721	GAGGGGTCCAAATGGGGCGGAGCTPACAAAGGGGACCTTTCTGAGCGAGGGGACCAACCTGGG	780
QY	781	ATGCTGACCTCTGTGCAAAACGGTCACTTTGGAGTCCAGACAGACAGACCCCACTGA	834
Db	781	ATGCTGACCTCTGTGCAAAACGGTCACTTTGGAGTCCAGACAGACAGACCCCACTGA	834
RESULT 6			
AF237290			
LOCUS	AF237290	1022 bp	mRNA linear PRI 02-AUG-2001
DEFINITION	Homo sapiens uridine-cytidine kinase 1 (UCK1) mRNA, complete cde.		
ACCESSION	AF237290		
VERSION	AF237290.1	GI:13506764	
KEYWORDS			
SOURCE			
ORGANISM	Homo sapiens (human)		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo. 1 (bases 1 to 1022)		
AUTHORS	Van Rompay,A.R., Norda,A., Linden,K., Johansson,M. and Karlsson,A. Phosphorylation of uridine and cytidine nucleoside analogs by two human uridine-cytidine kinases		
JOURNAL	Mol. Pharmacol. 59 (5), 1181-1186 (2001)		
MEDLINE	21203813		
PUBMED	11306702		
REFERENCE	2 (bases 1 to 1022)		
AUTHORS	Van Rompay,A.R., Linden,K., Norda,A., Zhu,C., Zheng,X., Johansson,M. and Karlsson,A.		
TITLE	Human uridine-cytidine kinase 1 and 2 : rate limiting enzymes		
JOURNAL	Unpublished		
REFERENCE	3 (bases 1 to 1022)		

AUTHORS			
Van Rompay, A. R., Lindén, K., Norda, A., Zhu, C., Zheng, X., Johansson, M. and Karlsson, A.			
TITLE			
Direct Submission			
JOURNAL			
Submitted (19-FEB-2000) IMPI, Clinical Virology, Huddinge University Hospital, Stockholm 14186, Sweden			
FEATURES			
SOURCE			
1..1022			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
1..1022			
/gene="UCLK1"			
5..838			
/gene="UCLK1"			
/function="(phosphorylates uridine and cytidine to uridine monophosphate and cytidine monophosphate)"			
/codon_start=1			
/product="uridine-cytidine kinase 1"			
/protein_id="AAK8324.1"			
/db_xref="GI:13506765"			
/translation="MASAGDEDESPAEADRPQRPDLGVSGTASGSTVCEKIM ELGQNEVEORQRKVLISQDRPYKVLTAQKAKALGQVNFDPDAFNDLMRLTK NIVGATVEPVYDFVTHSRLEPTVVYPADVLFEGILVEYSOEDIMFLRFPVDT DSVRLSRVLRPDRRDRLEOLITQYTTTPVKAFAEERLCPTKVAVLIIPIGVDMWV AIIIVIHIDILINDICIKMRGSSNRSYKRTFSEPDHPRMLTSGKSHLESSRP H"			
BASE COUNT			
227 a 282 c 324 g 189 t			
ORIGIN			
Query Match 93.9%; Score 783; DB 9; Length 1022;			
Best Local Similarity 99.9%; Pred. No. 0;			
Matches 833; Conservative 0; Mismatches 1; Indels 0; Gaps 0;			
QY	1	ATGGCTTCGGGGAGGCGAAGACTCGGAGAGCCCGCGCGAGCCGACCGTCCGCAC	60
DB	5	ATGGCTTCGGGGAGGCGAAGACTCGGAGAGCCCGCGCGAGCCGACCGTCCGCAC	64
QY	61	CAGCGGCCCTTCTCGATAGGGGTGAGCGCGGCACTGCCAGCGGAAATCGACCGTGT	120
DB	65	CAGCGGCCCTTCTCGATAGGGGTGAGCGCGGCACTGCCAGCGGAAATCGACCGTGT	124
QY	121	GAGGAAGATCATGAGATTCTGTGGACAAGACGAGGTGAAACAGCGGCAGCGAAGCTGTC	180
DB	125	GAGGAAGATCATGAGATTCTGTGGACAAGACGAGGTGAAACAGCGGCAGCGAAGCTGTC	184
QY	181	ATCCMGAGCCAGGACAGGTTCTACAAGGTCCTGACGGCAGACGAAAGGCCCAAGCCCTG	240
DB	185	ATCCMGAGCCAGGACAGGTTCTACAAGGTCCTGACGGCAGACGAAAGGCCCAAGCCCTG	244
QY	241	AAAGGACAGTCAATTTTGGACATCCAGATGCCCTTGTATATGATTTGATGACAGGACT	300
DB	245	AAAGGACAGTCAATTTTGGACATCCAGATGCCCTTGTATATGATTTGATGACAGGACT	304
QY	301	CTGAAGAACATCGTGAAGGCGAAACGCTGAGGTCGACCTATGATTTTGTGACACAC	360
DB	305	CTGAAGAACATCGTGAAGGCGAAACGCTGAGGTCGACCTATGATTTTGTGACACAC	364
QY	361	TCAAGGTTACACGAGACCAAGGTGCTACCCCTGCGGACGAGTTCGTTGAGGGGATC	420
DB	365	TCAAGGTTACACGAGACCAAGGTGCTACCCCTGCGGACGAGTTCGTTGAGGGGATC	424
QY	421	TTGGTGTTCACAGCCAGAGATCCGGGACATGTTCCACCTGCCTCTTGTGAGACAC	480
DB	425	TTGGTGTTCACAGCCAGAGATCCGGGACATGTTCCACCTGCCTCTTGTGAGACAC	484
QY	481	GACTCCGACGTGAGGCTGTCTCGAAGATTCCTCGGACGTCGCGCCGAGGAGGACCTG	540
DB	485	GACTCCGACGTGAGGCTGTCTCGAAGATTCCTCGGACGTCGCGCCGAGGAGGACCTG	544
QY	541	GAGGAGATTCTGAGGACGTACACACTTGTGTGAAAGCGCGCTTCGAGGAGTTCGTGCTG	600
DB	545	GAGGAGATTCTGAGGACGTACACACTTGTGTGAAAGCGCGCTTCGAGGAGTTCGTGCTG	604

```

Qy 601 CCGACAAAGAGTAGTCCGATGTGATCATCCACAGAGAGTGACAATATGTTGCCATC 660
Db 605 CCGACAAAGAGTAGTCCGATGTGATCATCCGAGAGAGTGACAAATATGTTGCCATC 664
Qy 661 AACCTGATCGTGCACACATCCAGACATTTCTGAATGTGACATCTGCAATGACACCGA 720
Db 665 AACCTGATCGTGCACACATCCAGACATTTCTGAATGTGACATCTGCAATGACACCGA 724
Qy 721 GGAGGATCTCAATGGGCGGAGCTTCTTCTGAGCCAGGGGACACCCCTGGG 780
Db 725 GGAGGATCTCAATGGGCGGAGCTTCTTCTGAGCCAGGGGACACCCCTGGG 784
Qy 781 ATGTGACCTCTGCAACCGTCACTTTGGAGTCCACAGACAGACCCCACTGA 834
Db 785 ATGTGACCTCTGCAACCGTCACTTTGGAGTCCACAGACAGACCCCACTGA 838

RESULT 7
LOCUS BD146824 753 bp DNA linear PAT 17-JAN-2003
DEFINITION Primer for synthesizing full-length cDNA and use thereof.
ACCESSION BD146824
VERSION BD146824.1 GI:27852582
KEYWORDS JP 2002191363-A/1667.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 753)
Oca,T., Isogai,T., Nishikawa,T., Hayashi,K., Saio,K., Yamamoto,J.,
Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
Primer for synthesizing full-length cDNA and use thereof
Patent: JP 2002191363-A 1667 09-JUL-2002;
HELIX RESEARCH INSTITUTE
OS Homo sapiens (human)
PN JP 2002191363-A/1667
PD 09-JUL-2002
PF 28-JUL-2000 JP 2000280990
PI TOSHIO OTA,TAKAO ISOGAI,TETSUO NISHIKAWA,KOJI HAYASHI,KAORU
PI SAITO,
PI JUNICHI YAMAMOTO,SHIZUKO ISHII,TOMOYASU SUGIYAMA,AI WAKAMATSU,
PI KEIICHI NAGAI,TETSUJI OTSUKI
PC: C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/
10,
PC C12P21/02,C12P1/68/C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof FH Key
FT source 1..753
Location/Qualifiers
1..753
/organism="Homo sapiens (human)"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

BASE COUNT 158 a 196 c 261 g 135 t 3 others

ORIGIN
Query Match 68.9%; Score 575; DB 6; Length 753;
Best Local Similarity 100.0%; Pred. No. 5.3e-292;
Matches 575; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGGCTTCGGGCGGAGCGAAGACTGCGAGAGCCCGCGCGGAGCGGACCGTCCGCAC 60
Db 95 ATGGCTTCGGGCGGAGCGAAGACTGCGAGAGCCCGCGCGGAGCGGACCGTCCGCAC 154
Qy 61 CAGCGGCCCTTCTCTGATAGGGGTGAGCGCGGCACTCCAGCGGAAAGTGCACCGTGTGT 120
Db 155 CAGCGGCCCTTCTCTGATAGGGGTGAGCGCGGCACTCCAGCGGAAAGTGCACCGTGTGT 214
Qy 121 GAGAAATCATGAGTGTCTGGACAGAACGAGGTGAAACGCGGACGCGGAAAGTGTGTC 180
Db 215 GAGAAATCATGAGTGTCTGGACAGAACGAGGTGAAACGCGGACGCGGAAAGTGTGTC 274

```

```

Qy 181 ATCTGAGCCAGGACAGGTTCTCAAGAGTCTCGACCGGACAGAGAGCCAGGCTTG 240
Db 275 ATCTGAGCCAGGACAGGTTCTCAAGAGTCTCGACCGGACAGAGAGCCAGGCTTG 334
Qy 241 AAAGACAGTACATTTTGAACCATCCAGATCCCTTGTATGATTTGATGACAGAGCT 300
Db 335 AAAGACAGTACATTTTGAACCATCCAGATCCCTTGTATGATTTGATGACAGAGCT 394
Qy 301 CTGAAGAACATCGTGAAGGGCAAAACGTTGAGAGTCCGACCTATGATTTTGGACAC 360
Db 395 CTGAAGAACATCGTGAAGGGCAAAACGTTGAGAGTCCGACCTATGATTTTGGACAC 454
Qy 361 TCAAGTTACACAGACACACGTTGTTACCTCGCGACGTTGTTGTTGAGGGCATC 420
Db 455 TCAAGTTACACAGACACACGTTGTTACCTCGCGACGTTGTTGTTGAGGGCATC 514
Qy 421 TTGTTGTTTACACGACGAGATCCGAGACATGTTCCACCTGCGCTTCTGTCGACACC 480
Db 515 TTGTTGTTTACACGACGAGATCCGAGACATGTTCCACCTGCGCTTCTGTCGACACC 574
Qy 481 GACTCCGACGTCAAGGCTGTCTCGAAGATTTCCGGAGCTGCGCGGAGGAGGACCTG 540
Db 575 GACTCCGACGTCAAGGCTGTCTCGAAGATTTCCGGAGCTGCGCGGAGGAGGACCTG 634
Qy 541 GAGGAGTTCTGACGCGAGTACACACCTTGTTGAA 575
Db 635 GAGGAGTTCTGACGCGAGTACACACCTTGTTGAA 669

RESULT 8
LOCUS AF125106 1395 bp mRNA linear PRI 07-FEB-2002
DEFINITION Homo sapiens uridine kinase mRNA, complete cds.
ACCESSION AF125106
VERSION AF125106.1 GI:18568108
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1395)
Kin,Y.R., Yu,L. and Zhao,S.Y.
Cloning of a new human cDNA similar to Mus musculus uridine kinase
mRNA
Unpublished
2 (bases 1 to 1395)
Ding,J.B., Yu,L. and Zhao,S.Y.
Direct Submission
Submitted (02-FEB-1999) Lab of Human Gene Research, Institute of
Genetics, Fudan University, No. 220 Handan Rd., Shanghai 200433,
People's Republic of China
Location/Qualifiers
1..1395
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
8..838
/codon_start=1
/product="uridine kinase"
/protein_id="AA175943.1"
/db_xref="GI:18568109"
/translation="MASAGEGARARAGANRPHORPFLIGVSGTASGKSTVCEKIME
LIGQNEVERNRKGVYILSDRFYKYLTAQKALKQYNRPHRPAFENDLNRITLKN
IVEGKVEVPTDFYTHSRLEPTVYVADVVLFGLIVFQELRDMFLRLPVDI
SDVRLSRVLRDVRGRDRLEQILTYVTFVKRPAFEFFCLPTKKYADVLI.PRGVDNVA
INLIYVHDIIDNDICIKWNRGSGNRTYKTFSEBGDHPMLTSGSKRSLSSSRPH
"

BASE COUNT 325 a 352 c 411 g 307 t

ORIGIN
Query Match 62.0%; Score 517; DB 9; Length 1395;
Best Local Similarity 99.8%; Pred. No. 2.3e-261;

```

Matches 567; Conservative 0; Mismatches 1; Indels 0; Gaps 0;				
QY	172	AGGTTGATCTCTGAGCCGAGACAGTTTACAAAGTCTTGAACGGAGAGCAAGGCC	231	
DB	176	AAAGTGATCTCTGAGCCGAGACAGTTTACAAAGTCTTGAACGGAGAGCAAGGCC	235	
QY	232	AAAGCCTTGAAGAGACAGTCAATTTTGAACCATTCAGATCCCTTGAATATGATTTGATG	291	
DB	236	AAAGCCTTGAAGAGACAGTCAATTTTGAACCATTCAGATCCCTTGAATATGATTTGATG	295	
QY	292	CACAGACCTCTGAAGAACATCTGTGAAGGGCAAAACGGTGAAGGTGCCCATATGATTTT	351	
DB	296	CACAGACCTCTGAAGAACATCTGTGAAGGGCAAAACGGTGAAGGTGCCCATATGATTTT	355	
QY	352	GTGACACACTCAAGTTTACAGAGACCAAGTGTCTACCCCTGCGACGTTGTTCTGTTT	411	
DB	356	GTGACACACTCAAGTTTACAGAGACCAAGTGTCTACCCCTGCGACGTTGTTCTGTTT	415	
QY	412	GAGGGATCTTGGTGTCTACAGCCAGAGATCCGGGACATGTTCACCTGCCCTCTTC	471	
DB	416	GAGGGATCTTGGTGTCTACAGCCAGAGATCCGGGACATGTTCACCTGCCCTCTTC	475	
QY	472	GTGACACACGACTCCGACGTGCTGTCTGAAAGTTCTCCGGACGTGCGCCGAGGG	531	
DB	476	GTGACACACGACTCCGACGTGCTGTCTGAAAGTTCTCCGGACGTGCGCGGGGG	535	
QY	532	AGGGACCTGGAAGAGATTCTGACGAGTACACACCTTCTGTGAAGCCGGCTTCAGAGAG	591	
DB	536	AGGGACCTGGAAGAGATTCTGACGAGTACACACCTTCTGTGAAGCCGGCTTCAGAGAG	595	
QY	592	TTTGTGCTGCGCAAAAGATATGCGATGTATCATCCACAGAGAGTGAATATG	651	
DB	596	TTTGTGCTGCGCAAAAGATATGCGATGTATCATCCACAGAGAGTGAATATG	655	
QY	652	GTTCGATCACTGATCTGTGACGACATTCAGAGACATTTGATGATGATCTGCAAA	711	
DB	656	GTTCGATCACTGATCTGTGACGACATTCAGAGACATTTGATGATGATCTGCAAA	715	
QY	712	TGGCACCGAGAGAGTTCATGGCGGGA	739	
DB	716	TGGCACCGAGAGAGTTCATGGCGGGA	743	
RESULT 9				
BC015547				
LOCUS				
DEFINITION				
ACCESSION				
VERSION				
KEYWORDS				
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REMARK				
COMMENT				

R. M.				
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/BLND at: http://image.llnl.gov				
Series: IRK Plate: 20 Row: 5 Column: 5				
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 13899252.				
FEATURES				
source				
1..2072				
/organism="Homo sapiens"				
/mol_type="mRNA"				
/db_xref="taxon:9606"				
/clone="MGC:9668 IMAGE:3845821"				
/tissue_type="Colon, adenocarcinoma"				
/clone_id="NIH_MGC_65"				
/lab_host="DH10B"				
/note="Vector: PCMV-SPORT6"				
87..692				
/codon_start=1				
/product="Similar to uridine-cytidine kinase 1"				
/protein_id="AAH15547.1"				
/db_xref="GI:15930230"				
/translation="MASAGEDCESPAPRDRPHQRPFLIGVSGTASGKTVCEKIM				
ELIGONEVEORQRKTVILSDRFYVLTAEQAKALKQYNDHPDAPNDLHRTLK				
NIVEGTVEVPTVDFTVSHRLPBTIVVADVVLEGLIVFSQEIADWFHRLFVDT				
DSDVLSRDKVPCDHPTRSGYCHQPDAAHPHSEW"				
BASE COUNT				
455 a 559 c 637 g 421 t				
ORIGIN				
Query Match				
Best Local Similarity 100.0%; Pred. No. 1.3e-25; Length 2072;				
Matches 508; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
QY	1	ATGCTTGGCGGAGGCGAAGACTGCGAGACCCCGCGAGGCCGACCGTCCGAC	60	
DB	87	ATGCTTGGCGGAGGCGAAGACTGCGAGACCCCGCGAGGCCGACCGTCCGAC	146	
QY	61	CAGCGGCTTCTCTGATAGGGGTGAGCGGCGGCACTGCGAGGGGAAGTGCACCGTGT	120	
DB	147	CAGCGGCTTCTCTGATAGGGGTGAGCGGCGGCACTGCGAGGGGAAGTGCACCGTGT	206	
QY	121	GAGAAGATCATGAGTGTCTGAGGACGAGAGAGTGAACAGCGGACGGAAGGTGTC	180	
DB	207	GAGAAGATCATGAGTGTCTGAGGACGAGAGAGTGAACAGCGGACGGAAGGTGTC	266	
QY	181	ATCTGAGCCAGAGACGTTCTTACAAAGTCTGACGCGACAGACGAAAGGCGCTTG	240	
DB	267	ATCTGAGCCAGAGACGTTCTTACAAAGTCTGACGCGACAGACGAAAGGCGCTTG	326	
QY	241	AAAGACATTAATTTTACATCCAGATGCTTGTGATTAATGATTTGATGACAGGACT	300	
DB	327	AAAGACATTAATTTTACATCCAGATGCTTGTGATTAATGATTTGATGACAGGACT	386	
QY	301	CTGAAGAACATGTTGAGGAGGCAAAACGGTGAAGTGCAGCTATGATTTTGAACAC	360	
DB	387	CTGAAGAACATGTTGAGGAGGCAAAACGGTGAAGTGCAGCTATGATTTTGAACAC	446	
QY	361	TCAAGTTTACAGAGACACGAGTGTCTTACCTGCGGACGTTCTGTTTGAAGGACATC	420	
DB	447	TCAAGTTTACAGAGACACGAGTGTCTTACCTGCGGACGTTCTGTTTGAAGGACATC	506	
QY	421	TTGGTGTCTTACAGCAGAGAGATCCGGGACATGTTCCACTGCGCTCTTGTGACAC	480	
DB	507	TTGGTGTCTTACAGCAGAGAGATCCGGGACATGTTCCACTGCGCTCTTGTGACAC	566	
QY	481	GACTCGACGTGAGGCTGTCTGAAAGAG	508	
DB	567	GACTCGACGTGAGGCTGTCTGAAAGAG	594	
RESULT 10				
AK057848				
LOCUS				
2228 bp				
mRNA				
linear				
PRI 31-OCT-2001				

DEFINITION Homo sapiens cDNA FLJ25119 fis, clone CBR05878, highly similar to
UNIDISE KINASE (EC 2.7.1.48).
ACCESSION AK057848.1 GI:16553809
VERSION AK057848.1
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
AUTHORS Tashiro,H., Yamazaki,M., Watanabe,K., Kumagai,A., Itakura,S.,
Fukuzumi,Y., Fujimori,Y., Komiya,M., Suzuki,Y., Hata,H.,
Nakagawa,K., Mizuno,S., Morinaga,M., Kawamura,T., Sugiyama,T.,
Irie,R., Otsuki,T., Sato,H., Nishikawa,T., Sugiyama,A.,
Kawakami,B., Nagai,K., Isogai,T. and Sugano,S.
TITLE NEDO human cDNA sequencing project
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2228)
AUTHORS Sugano,S. and Suzuki,Y.
TITLE Direct Submission
JOURNAL Submitted (24-OCT-2001) Sumio Sugano, Institute of Medical Science,
University of Tokyo, Laboratory of Genome Structure, Human Genome
Center; Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan
(E-mail:cdna@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
Fax:81-3-5449-5416)
COMMENT NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing;
Research Association for Biotechnology (RAB); cDNA library
construction and 5'-end one pass sequencing; Institute of Medical
Science, University of Tokyo, Laboratory of Genome Structure, Human
Genome Center; 3'-end one pass sequencing; RAB; clone selection for
full insert sequencing; RAB and Helix Research Institute.
FEATURES
source
location/Qualifiers
1..2228
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CBR05878"
/tissue_type="brain"
/clone_id="CBR"
/note="Cloning vector: pME18SFL3"
BASE COUNT 478 a 598 c 696 g 456 t
ORIGIN
Query Match 59.6%; Score 497; DB 9; Length 2228;
Best Local Similarity 100.0%; Pred. No. 8.1e-251;
Matches 497; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 107 AGTCGACCGTGTGTGAGAAGATCATGAGTTGCTGGGACAGAAAGAGTGAACAGCGGC 166
DB 212 AGTGAACGCTGTGTGAGAAGATCATGAGTTGCTGGGACAGAAAGAGTGAACAGCGGC 271
QY 167 AGCGGAAGGTGTGATCTCTGAGCCAGGACAGGTTCTCAAGAGTCTTGAACGCGCAGAGAGA 226
DB 272 AGCGGAAGGTGTGATCTCTGAGCCAGGACAGGTTCTCAAGAGTCTTGAACGCGCAGAGAGA 331
QY 227 AGGCGAAGGCTTGAAGGACAGTACAAATTTGACATCCAGAGCTTGTGATATATAT 286
DB 332 AGGCGAAGGCTTGAAGGACAGTACAAATTTGACATCCAGAGCTTGTGATATATAT 391
QY 287 TGATGACAGGACTCTGAGAAACATGCTGAGGCAAAACGCTGAGAGTCCGACCTATG 346
DB 392 TGATGACAGGACTCTGAGAAACATGCTGAGGCAAAACGCTGAGAGTCCGACCTATG 451
QY 347 ATTTTGTGACACTCAAGGTTACAGAGACCAAGTGTCTACCTTGGCGAGCTGTTC 406
DB 452 ATTTTGTGACACTCAAGGTTACAGAGACCAAGTGTCTACCTTGGCGAGCTGTTC 511
QY 407 TGTGTGAGGGGATCTGTGTCTACAGCCAGAGGATCCGGGAGATGTTCACTCTGCGC 466
DB 512 TGTGTGAGGGGATCTGTGTCTACAGCCAGAGGATCCGGGAGATGTTCACTCTGCGC 571
QY 467 TCTTGTGACACCGACTCCGACGTCAAGCTGTCTCGAAGAGTTCTCCGGAGAGTGTGCC 526

DB 572 TCTTGTGACACCGACTCCGACGTCAAGCTGTCTGAGAGTTCTCCGGAGAGTGTGCC 631
QY 527 GAGGAGAGGACCTTGAAGCAGATTTCTGACGATACACACCTTGTGTAAGCCGCTTCG 586
DB 632 GAGGAGAGGACCTTGAAGCAGATTTCTGACGATACACACCTTGTGTAAGCCGCTTCG 691
QY 587 AGGAGTTCTGCGCTGCGC 603
DB 692 AGGAGTTCTGCGCTGCGC 708
RESULT 11
AX540411 734 bp DNA linear PAT 23-NOV-2002
LOCUS AX540411
DEFINITION Sequence 23 from Patent WO02055738.
ACCESSION AX540411
VERSION AX540411.1 GI:25273437
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
AUTHORS Panzer,S.R., Lincoln,S.E., Altus,C.M., Dufour,G.E., Hillman,D.L.,
Jones,A.L., Lam,T.C., Liu,T.F., Harris,B., Flores,V., Daffo,A.,
Marwaha,R., Chen,A.U., Chang,S.C., Gerstlin,E.H., Petralia,C.H.,
David,M.H. and Lewis,S.A.
TITLE Molecules for disease detection and treatment
JOURNAL Patent: WO 02055738-A 23 18-Jul-2002;
Incyte Genomics, Inc. (US)
FEATURES
source
location/Qualifiers
1..734
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/note="Incyte ID No: LI:235557.12:2001JAN12"
BASE COUNT 166 a 181 c 226 g 154 t 7 others
ORIGIN
Query Match 40.5%; Score 338; DB 6; Length 734;
Best Local Similarity 100.0%; Pred. No. 8.6e-167;
Matches 338; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 266 CAGATGCTTGTGATATGATTTGATGACAGGACTCTGAAGAACATCGTGGAGGGA 325
DB 12 CAGATGCTTGTGATATGATTTGATGACAGGACTCTGAAGAACATCGTGGAGGGA 71
QY 326 CGGTGAGAGTGCAGACCTATGATTTTGTGACACACTCAAGTTTACAGAGACCAAGCTGG 385
DB 72 CGGTGAGAGTGCAGACCTATGATTTTGTGACACACTCAAGTTTACAGAGACCAAGCTGG 131
QY 386 TCTACCTTGGGAGCGTGTCTGTTTGAAGGACATCTTGTGTCTTACAGCCAGAGATCC 445
DB 132 TCTACCTTGGGAGCGTGTCTGTTTGAAGGACATCTTGTGTCTTACAGCCAGAGATCC 191
QY 446 GGGACATGTTCCACCTGCGCTCTTGTGAGACACGACCTCCAGCGTCAAGCTGTCTGAA 505
DB 192 GGGACATGTTCCACCTGCGCTCTTGTGAGACACGACCTCCAGCGTCAAGCTGTCTGAA 251
QY 506 GAGTTCTCCGGGACGTGCGCCGAGAGGAGGACCTGAGAGAGATTCTGAGCAGTACCA 565
DB 252 GAGTTCTCCGGGACGTGCGCCGAGAGGAGGACCTGAGAGAGATTCTGAGCAGTACCA 311
QY 566 CTTTGTGAGAGCGGCTTTCAGAGAGTTTCTGCTGCG 603
DB 312 CTTTGTGAGAGCGGCTTTCAGAGAGTTTCTGCTGCG 349
RESULT 12
AL358781 147492 bp DNA linear PRI 06-OCT-2001
LOCUS AL358781/c
DEFINITION Human DNA sequence from clone RP11-334J6 on chromosome 9, complete

sequence.
 AL358781
 AL358781.19 GI:13751418
 HTG.
 Homo sapiens (human)
 SOURCE
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 147492)
 Corby,N.
 Direct Submission
 Submitted (06-OCT-2001) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
 requests: clonerequest@sanger.ac.uk
 On April 21, 2001 this sequence version replaced gi:13395549.
 During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.
 This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats: all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em; EMBL; SW; SWISSPROT; Tr; TREMBL; Wp; WORMPEP; Information on the WORMPEP database can be found at
 http://www.sanger.ac.uk/Projects/C_elegans/wormpep
 This sequence was generated from part of bacterial clone contigs of human chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping Group. Further information can be found at
 http://www.sanger.ac.uk/HGP/Chr9
 RP11-33436 is from the library RPCT-11.2 constructed by the group of Pieter de Jong. For further details see
 http://www.chori.org/bacpac/home.htm
 VECTOR: pBACe3.6
 This sequence is the entire insert of clone RP11-33436 The true left end of clone RP11-4087 is at 113870 in this sequence. The true right end of clone RP11-643B14 is at 63282 in this sequence.
 FEATURES
 source
 1. 147492
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="9"
 /clone="RP11-33436"
 /clone_1lb="RPCT-11.2"
 30143. 30199
 /note="Sequence from overlapping clone RP11-643B14 (AL354855). Assembly confirmed by restriction digest."
 misc_feature
 BASE COUNT 32592 a 36290 c 37577 g 40633 t
 ORIGIN
 Query Match 21.9%; Score 183; DB 9; Length 147492;
 Best Local Similarity 100.0%; Pred. No. 4.8e-85;
 Matches 183; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 652 GTTGCCATCACTGATCGTCGACGACATTCAGAGCACTTCTGAATGCTGCACTTGCAAA 711
 Db 129488 GTTGCCATCACTGATCGTCGACGACATTCAGAGCACTTCTGAATGCTGCACTTGCAAA 129429
 QY 712 TGGCACCAGAGAGGCTCAATGGCGAGATCAAGGAGCACTTTCTGACCGAGGGAC 771
 Db 129428 TGGCACCAGAGAGGCTCAATGGCGAGATCAAGGAGCACTTTCTGACCGAGGGAC 129369
 QY 772 CACCTTGAGATGCTGACCTTGGCAACGGTCACATTGGAGTCCAGACGACACCCAC 831
 Db 129368 CACCTTGAGATGCTGACCTTGGCAACGGTCACATTGGAGTCCAGACGACACCCAC 129309

QY 832 TGA 834
 Db 129308 TGA 129306
 RESULT 13
 BV077668/c
 LOCUS
 DEFINITION
 BV077668 660 bp DNA linear STS 31-MAY-2003
 S212P6227FD8.T0 CZECHII/E1 Mus musculus STS genomic, sequence
 tagged site.
 BV077668
 BV077668.1 GI:31193463
 STS.
 Mus musculus (house mouse)
 SOURCE
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 660)
 Lander,E.S., Lindblad-Toh,K. and Daly,M.J.
 The mosaic structure of variation in the laboratory mouse genome
 Nature 420 (6915), 574-578 (2002)
 22354684
 12468852
 COMMENT
 Contact: Kerstin Lindblad-Toh
 Whitehead Institute for Biomedical Research, Center for Genome
 Research
 320 Charles Street, Cambridge, MA 02141, USA
 Tel: 6172521477
 Fax: 6172580903
 Email: kersti@genome.wi.mit.edu
 Primer A: None
 Primer B: None
 STS size: 660
 STS protocol:
 WGS-discovery: Paired-end low-coverage whole genome shotgun reads
 were generated from 12951/SVIMJ, C3H/HeJ, and BALB/cByJ. The WGS
 reads were placed uniquely on the MGS5C3 C57BL/6J assembly and SNP
 detection was carried out by SSAHA-SNP. 225,000 reads were
 annotated
 as STSs and 81,000 SNPs were annotated with alleles from C57BL/6J
 and the strain from which the particular read came. The validation
 rate for these SNPs was estimated at approximately 98%.
 FEATURES
 source
 1. 660
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /strain="CZECHII/E1"
 /db_xref="taxon:10090"
 /map="2 22-615 32611326-32611918"
 /clone_1lb="CZECHII/E1"
 <1. .>660
 STS
 BASE COUNT 174 a 196 c 174 g 116 t
 ORIGIN
 Query Match 4.3%; Score 36; DB 11; Length 660;
 Best Local Similarity 100.0%; Pred. No. 3e-07;
 Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 435 CCAGAGATCCGGAGATGTCACCTGCGCCTCT 470
 Db 435 CCAGAGATCCGGAGATGTCACCTGCGCCTCT 400
 RESULT 14
 MUSURKI
 LOCUS
 DEFINITION
 Mus musculus uridine kinase mRNA, partial cds.
 ACCESSION
 L31783
 VERSION
 L31783.1 GI:471980
 KEYWORDS

SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS 1 (bases 1 to 1810)
TITLE The functions and consensus motifs of nine types of peptide segments that form different types of nucleotide-binding sites
JOURNAL Eur. J. Biochem. 222 (1), 9-19 (1994)
MEDLINE 94259063
PUBMED 8200357
REFERENCE 2 (bases 1 to 1810)
AUTHORS Kopp, P.A. and Traut, T.W.
TITLE Cloning and expression of a cDNA encoding uridine kinase from mouse brain
JOURNAL Arch. Biochem. Biophys. 336 (1), 105-112 (1996)
MEDLINE 97108719
PUBMED 8951040
FEATURES
source
Location/Qualifiers
1..1810
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/tissue_type="Brain"
/EC_number="2.7.1.48"
/standard_name="uridine-cytidine kinase"
/codon_start=1
/product="uridine kinase"
/protein_id="AA50568.1"
/db_xref="GI:471981"
/translation="RQPRPFLIGVSGTASGKSTVCEKIMELLGQNEVDRORKLVI
LSQCFKVLTAQKAKAKQYFNDFDPAFNDMLKTKNIYEGKTVFPTDYDT
HSRLPRTVVYPADVLEGLIVYTOEIRDMFLRLPVDPSVRLSRRLDVGK
RDLEQILQVYAPVKAPEERECPTKTKADYIRPGVDNVAIYLQIHQDILNGDL
CKRRGCPNGNHRKTPPEPDHFGVLTGRSHLESSSRP"
misc_binding
join(37..63,412..426,451..465)
/bound_moiety="ATP at catalytic site"
polya_signal
1790..1795
BASE COUNT 411 a 482 c 512 g 405 t
ORIGIN
Query Match 4.3%; Score 36; DB 10; Length 1810;
Best Local Similarity 100.0%; Pred. No. 2.8e-07;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

435 CCAGGAGATCCGGACATGTCACCTGCGCCTTT 470
Db 384 CCAGGAGATCCGGACATGTCACCTGCGCCTTT 419

RESULT 15
LOCUS BC025146 1959 bp mRNA linear ROD 16-Apr-2003
DEFINITION Mus musculus uridine monophosphate kinase, mRNA (cDNA clone
MGC:36231 IMAGE:4913412), complete cds.
ACCESSION BC025146
VERSION BC025146.1 GI:19263563
KEYWORDS MGC.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS 1 (bases 1 to 1959)
TITLE Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heien, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Schetz, T.E., Brownstein, M.J., Usadi, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Rana, S.S., Loquellano, N.A., Peters, G.J.,
Adams, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J.,

McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Morley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Raney, J.J., Helton, E., Ketterman, M., Madan, A., Rodrigues, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalske, U., Smalls, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
22388257
12477932
REFERENCE 2 (bases 1 to 1959)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (05-MAR-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
COMMENT Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: amg@bcm.tmc.edu
Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
A.N., Gibbs, R.A.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LIN at: <http://image.llnl.gov>
Series: IRAC Plate: 61 Row: P Column: 19
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis, similarity but not identity to protein.
Location/Qualifiers
1..1959
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="MGC:36231 IMAGE:4913412"
/tissue_type="Salivary gland, 10 week old female mouse"
/clone_lib="NCI CGAP_SG2"
/lab_host="DH10B"
/note="Vector: PCMV-SPORT6"
1..1959
/gene="Umpk"
/db_xref="LocusID:22245"
/db_xref="MGI:98904"
57..890
/codon_start=1
/product="Umpk protein"
/protein_id="AAH25146.1"
/db_xref="GI:19263564"
/db_xref="LocusID:22245"
/translation="MASAGGSGSBSAAPADRPQRPPLIGVSGTASGKSTVCEKIM
ELLGQNEVDRORKLVIISQCFKVLTAQKAKQYFNDFDPAFNDMLKTKNIYEGKTVFPTDYDT
HSRLPRTVVYPADVLEGLIVYTOEIRDMFLRLPVDPSVRLSRRLDVGKRDLEQILQVYAPVKAPEERECPTKTKADYIRPGVDNVAIYLQIHQDILNGDLCKRRGCPNGNHRKTPPEPDHFGVLTGRSHLESSSRP"
H"
BASE COUNT 466 a 511 c 561 g 421 t
ORIGIN
Query Match 4.3%; Score 36; DB 10; Length 1959;

Best Local Similarity 100.0%; Pred. No. 2.8e-07;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 435 CCAGGAGATCCGGACATGTTCCACCTGCGCCTCTT 470
|||
491 CCAGGAGATCCGGACATGTTCCACCTGCGCCTCTT 526

Search completed: November 25, 2003, 05:12:22
Job time : 3245.82 secs

This Page Blank (uspto)


```

Pred. No.: 1.26e-37 Length: 277
Score: 45.00 Matches: 45
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 16.25% Indels: 0
DB: 1 Gaps: 0

US-09-896-522-3 (1-834) x UCK1_MOUSE (1-277)

QY 298 ACCTGTGAAGACATCGTGGAGGGGAAACGCTGAGGTGCCGACCTTATGATTTTGACACA 357
    |||
Db 100 ThrlenuysasnlleValgluglyLythrValglValProtnrTyraepPhevalThr 119

QY 358 CACTCAAGTTTACAGACGACGAGTGTCTACCCCTCGAGACGTTCTGTTTGAGAGGC 417
    |||
Db 120 HisSerArgLeuProGluTrnValValTyProAlaPValValLeuPheGly 139

QY 418 ATCTGTGTCTTCTAC 432
    |||
Db 140 IleLeuValPheTy 144

RESULT 3
UCK2_HUMAN
ID UCK2_HUMAN STANDARD; PRT; 261 AA.
AC Q9BZX2; Q96KG5;
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last sequence update)
DE Uridine-cytidine kinase 2 (EC 2.7.1.48) (UCK 2) (Uridine
GN monophosphokinase 2) (Cytidine monophosphokinase 2).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. AND CHARACTERIZATION.
RX MEDLINE=21203813; PubMed=11306702;
RA Van Rompay A.R., Norda A., Linden K., Johanson M., Karlsson A.;
RT "Phosphorylation of uridine and cytidine nucleoside analogs by two
RL human uridine-cytidine kinases."
RM Mol. Pharmacol. 59:1181-1186(2001).
RN [2]
RP SEQUENCE OF 15-261 FROM N.A.
RX TISSUE=fibrosarcoma;
RA MEDLINE=21385121; PubMed=11494055;
RA Koizumi K., Shimamoto Y., Azuma A., Mataya Y., Matsuda A., Sasaki T.,
RT "Cloning and expression of uridine/cytidine kinase cDNA from human
RL fibrosarcoma cells."
RN [3]
RP Int. J. Mol. Med. 8:273-278(2001).
CC -1- FUNCTION: Phosphorylates uridine and cytidine to uridine
CC monophosphate and cytidine monophosphate. Does not phosphorylate
CC deoxyribonucleosides or purine ribonucleosides. Can use ATP or GTP
CC as a phosphate donor. Can also phosphorylate cytidine and uridine
CC nucleoside analogs such as 6-azauridine, 5-fluorouridine, 4-
CC thioridine, 5-bromouridine, N(4)-acetylcytidine, N(4)-
CC benzoylcytidine, 5-fluorocytidine, 2-thiocytidine, 5-
CC methylcytidine, and N(4)-anisoylcytidine.
CC -1- CATALYTIC ACTIVITY: ATP + uridine = ADP + UMP.
CC -1- CATALYTIC ACTIVITY: ATP + cytidine = ADP + CMP.
CC -1- PATHWAY: Pyrimidine salvage pathway.
CC -1- TISSUE SPECIFICITY: Expressed in placenta.
CC -1- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

DR EMBL; AF236637; AAK14053.1; -.
DR EMBL; AB062451; BAB56162.1; -.
DR Genew; H0NC12562; UMPK.
DR InterPro; IPR006082; PRK.
DR InterPro; IPR006083; PRK URK.
DR InterPro; IPR00764; Uridine_kin.
DR Pfam; PF00485; PRK; 1.
DR PRINTS; PR00478; PHRIBLKINASE.
DR PRINTS; PR00988; URIDINKINASE.
DR TIGRFAMs; TIGR00235; udk; 1.
KW Transferase; Kinase; ATP-binding.
FT ND BIND 27 34 ATP (POTENTIAL).
SQ SEQUENCE 261 AA; 29299 MW; 71791346F091EBFD CRC64;

Alignment Scores:
Pred. No.: 4.59e-17 Length: 261
Score: 25.00 Matches: 25
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 9.03% Indels: 0
DB: 1 Gaps: 0

US-09-896-522-3 (1-834) x UCK2_HUMAN (1-261)

QY 565 ACCTGTGAAGCGCGCTTGGAGAGTCTGCGTCCGACAAAGATGCGCATG 624
    |||
Db 187 TrnPhValValProAlaPheGluGluPheCysLeuProTrnValValValVal 206

QY 625 ATCATCCACGAGGA 639
    |||
Db 207 IleIleProArgGly 211

RESULT 4
UCK2_MOUSE
ID UCK2_MOUSE STANDARD; PRT; 261 AA.
AC Q99PM9;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last sequence update)
DE Uridine-cytidine kinase 2 (EC 2.7.1.48) (UCK 2) (Uridine
GN monophosphokinase 2) (Cytidine monophosphokinase 2).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21203813; PubMed=11306702;
RA Van Rompay A.R., Norda A., Linden K., Johanson M., Karlsson A.;
RT "Phosphorylation of uridine and cytidine nucleoside analogs by two
RL human uridine-cytidine kinases."
RM Mol. Pharmacol. 59:1181-1186(2001).
CC -1- FUNCTION: Phosphorylates uridine and cytidine to uridine
CC monophosphate and cytidine monophosphate. Does not phosphorylate
CC deoxyribonucleosides or purine ribonucleosides. Can use ATP or GTP
CC as a phosphate donor. Can also phosphorylate cytidine and uridine
CC nucleoside analogs such as 6-azauridine, 5-fluorouridine, 4-
CC thioridine, 5-bromouridine, N(4)-acetylcytidine, N(4)-
CC benzoylcytidine, 5-fluorocytidine, 2-thiocytidine, 5-
CC methylcytidine, and N(4)-anisoylcytidine (By similarity).
CC -1- CATALYTIC ACTIVITY: ATP + uridine = ADP + UMP.
CC -1- CATALYTIC ACTIVITY: ATP + cytidine = ADP + CMP.
CC -1- PATHWAY: Pyrimidine salvage pathway.
CC -1- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

CC -----
DR EMBL: AF236636; AAK14052.1; -.
DR MGD; MGI:1931744; UCK2.
DR InterPro; IPR006082; PRK.
DR InterPro; IPR006083; PRK URK.
DR InterPro; IPR00764; Uridine_kin.
DR Pfam; PF00485; PRK; 1.
DR PRINTS; PR00478; PHRIBKINASE.
DR PRINTS; PR00988; URIDINKINASE.
DR TIGRFAMs; TIGR00235; uck; 1.
KM Transferase; Kinase; ATP-binding.
FT NP BIND 27 34 ATP (POTENTIAL).
SQ SEQUENCE 261 AA; 29404 MW; 780AA3CF5CA8155 CRC64;

Alignment Scores:
Pred. No.: 4,59e-17 Length: 261
Score: 25.00 Matches: 25
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 9.03% Indels: 0
DB: 1 Gaps: 0

US-09-896-522-3 (1-834) x UCK2_MOUSE (1-261)
QY 565 ACCTTCGAGCGCGCTTGGAGAGTTCCTGCGCAGACAGAGTATCGGATGTG 624
Db 187 ThreVallylProAlaValIleuPheGluGluPheCysLeuProTyrLysTyrAlaAspVal 206

QY 625 ATCATCCACGAGGA 639
Db 207 IleIleProArgLys 211

RESULT 5
ID UCK_DROME STANDARD; PRK; 260 AA.
AC Q9VC99;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Probable uridine-cytidine kinase (EC 2.7.1.48) (UCK) (Uridine
DE monophosphokinase) (Cytidine monophosphokinase).
GN CG6364.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC SMPAIN-Berkeley;
RX MEDLINE=20196006; PubMed=10731132.
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan X.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abell J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brockstein P., Brothier P.,
RA Butts K.C., Buam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foaier C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.U., Wei M.-H., Ibegwan C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kenton J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laeko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,

```

```

RA Liu X., Matrei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Mosnig A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Palenik K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
CC -1- CATALYTIC ACTIVITY: ATP + uridine = ADP + UMP.
CC -1- CATALYTIC ACTIVITY: ATP + cytidine = ADP + CMP.
CC -1- PATHWAY: Pyrimidine salvage pathway.
CC -1- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE003747; AAF56274.1; -.
DR FlyBase; FBgn0039179; CG6364.
DR InterPro; IPR006083; PRK URK.
DR InterPro; IPR00764; Uridine_kin.
DR Pfam; PF00485; PRK; 1.
DR PRINTS; PR00988; URIDINKINASE.
KM Hypothetical protein; Transferase; Kinase; ATP-binding.
FT NP BIND 34 41 ATP (POTENTIAL).
SQ SEQUENCE 260 AA; 29213 MW; 2EE57423704925B6 CRC64;

Alignment Scores:
Pred. No.: 8.73e-07 Length: 260
Score: 15.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.42% Indels: 0
DB: 1 Gaps: 0

US-09-896-522-3 (1-834) x UCK_DROME (1-260)
QY 388 TACCTCGAGCGAGCTGCTTGGAGGACATCTGTTCTTAC 432
Db 135 TyrProAlaAspValIleuPheGluGluLysIleValPheTyr 149

Search completed: November 25, 2003, 07:54:03
Job time : 10.6522 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: November 25, 2003, 07:50:54 ; Search time 8.99146 Seconds

(without alignments)

7849.063 Million cell updates/sec

Title: US-09-896-522-3

Perfect score: 277

Sequence: 1 atgagcttcgagcgagcgagcga.....ccagcagcagacaccacatga 834

Scoring table:

OLIGO	
Xgapop 60.0 , Xgapext 60.0	
Ygapop 60.0 , Ygapext 60.0	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 328717 seqs, 42310858 residues

Word size: 15

Total number of hits satisfying chosen parameters: 2

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODE=frame+n2p.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US09896522/rnucat_21112003.184146.3259/app_query.fasta.1.2830
-DB=Issued Patents AA -QFMT=fastaan -SUFFIX=n2p_ol1.ra1 -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human0.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=quality -THR MIN=15 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pro -NOR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09896522.QCEN_1.1.29@rnucat_21112003.184146.3259 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

Issued Patents AA.*
1: /cgn2_6/prodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/prodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/prodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/prodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/prodata/2/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/prodata/2/iaa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	260	93.9	260	4	US-09-536-647-2
2	45	16.2	277	4	US-09-536-647-3

ALIGNMENTS

RESULT 1
US-09-536-647-2
; Sequence 2, Application US/09536647
; Patent No. 6579708
; GENERAL INFORMATION:
; APPLICANT: Ho, Yen Sen

APPLICANT: Johnson, Randall
; TITLE OF INVENTION: No. 6579708e1 Human Uridine Kinase
; FILE REFERENCE: GP50020
; CURRENT APPLICATION NUMBER: US/09/536,647
; CURRENT FILING DATE: 2000-03-27
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 260
; TYPE: PRT
; ORGANISM: Human
US-09-536-647-2

Alignment Scores:

Pred. No.:	2,26e-244	Length:	260
Score:	260.00	Matches:	260
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	93.86%	Indels:	0
DB:	4	Gaps:	0

US-09-896-522-3 (1-834) x US-09-536-647-2 (1-260)

QY	1	ATGAGCTTCGAGCGAGCGAGAGAGCTGCGAGAGCCCGCGAGGCGACCGCTCGCAC	60
DB	1	MetAlaSerAlaGlyGlyGlnAspGlySerGlnProAlaProGlnAlaAspArgProHis	20
QY	61	CAGCGCCCTTCCTGATGAGGAGTGAACGGCGACACTGCCAGCGGAAGTGAACCGTGT	120
DB	21	GlnArgProPheLeuIleGlyAlaSerGlyThrAlaSerGlyLysSerThrValCys	40
QY	121	GAGAGATCATGAGTGTGCGGAGCAGACGAGTGGAGCAGCGGAGGAGTGTG	180
DB	41	GlnLysIleMetGlnLeuLeuGlnGlnAlaGlnGlnAlaGlnGlnArgGlnValVal	60
QY	181	ATCTCGAGCGCAGAGCAGGTTCTTCAAGGCTCTGACGCGCAGACAGAGCGCCTTG	240
DB	61	IleLeuSerGlnAspArgPheTyrLysValLeuThrAlaGlnGlnLysAlaLysAlaLeu	80
QY	241	AAAGCAGTACATTTTGAACCATCCAGATGCGCTTGTATGATGATGACAGAGACT	300
DB	81	LysGlyGlnTyrAsnPheAspHisProAlaPheAspAsnAspLeuMetHisArgThr	100
QY	301	CTGAAGATCATGCTGAGAGGCAAAACGCTGAGAGTCCGACCTTATGATTTTGACAC	360
DB	101	LeuLysAsnIleValGlnGlyLysThrValGlnValProThrTyrAspPheValThrHis	120
QY	361	TCAGGTTACAGAGCAGCGAGTGTCTACCTCGCGAGAGCTGTTCTGTTGAGGCGATC	420
DB	121	SerArgLeuProGlnThrThrValValTyrProAlaAspValValLeuPheGlnGlyIle	140
QY	421	TTGTGTCTTACAGCGCAGAGATCCGAGATGTTCCACTGCGCTCTTGTGACAC	480
DB	141	LeuValPheTyrSerGlnGlnIleArgAspMetPheIleuArgLeuPheValAspThr	160
QY	481	GACTCGAGTCAAGGCTGCTCGAAGAGTCTCCGGAGAGTCCGCCGAGGAGGACCTG	540
DB	161	AspSerAspValArgLeuSerArgValLeuArgAspValArgArgGlyLysArgAspLeu	180
QY	541	GAGCAGATCTCAGCAGATACACACACCTTCGTAAGCGCGCTTCGAGAGGTTGCGCTG	600
DB	181	GlnGlnIleLeuThrGlnTyrThrThrPheValLysProAlaPheGlnGlnLysCysLeu	200
QY	601	CCGACAAAGAGTATGCGAGTGTATCATCCAGAGAGTGGAGACATATGTTGTCATC	660
DB	201	ProThrLysLysTyrAlaAspValIleIleProArgGlyValAspAsnMetValAlaIle	220
QY	661	AACCTGATGTCACACATCCAGACATTTGATGAGATGAGATTTGCAAAAGGACCGA	720
DB	221	AsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysLysThrPheSarg	240
QY	721	GGAGGTCGAATGGCGAGCTACAGCGGACCTTTCTGAGCCAGGAGCAGCCTGAG	780

Db 241 GlyGlySerAenGlyArgSerTyrIlyArgThrPheSerGluProGlyAapHleProGly 260

RESULT 2

US-09-536-647-3
; Sequence 3, Application US/09536647
; Patent No. 6579708
; GENERAL INFORMATION:
; APPLICANT: Ho, Yen Sen
; APPLICANT: Johnson, Randall
; TITLE OF INVENTION: No. 6579708e1 Human Uridine Kinase
; FILE REFERENCE: GP50020
; CURRENT APPLICATION NUMBER: US/09/536,647
; CURRENT FILING DATE: 2000-03-27
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 277
; TYPE: PRT
; ORGANISM: Human
US-09-536-647-3

Alignment Scores:

Pred. No.:	2,21e-35	Length:	277
Score:	45.00	Matches:	45
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	16.25%	Indels:	0
DB:	4	Gaps:	0

US-09-896-522-3 (1-834) x US-09-536-647-3 (1-277)

QY	298	ACTGTGAAGACATCGTGAAGGCAAAACGGTGAAGTGCACCTATGATTTGTGACA	357
Db	100	ThreulysasnIleValGluGlyLysThrValGluValProtnTyrAspPheValThr	119
QY	358	CACCTAAGGTTACGACAGACACCGGTGCTTACCTCGCGACGTGTTGTTGAGGAC	417
Db	120	HisSerArgLeuProGluThrThrValValTyrProAlaSerValIleuPheGluGly	139
QY	418	ATCTTGTTGTTCTTAC	432
Db	140	IleLeuValPheTyr	144

Search completed: November 25, 2003, 08:01:01
Job time : 9.99146 secs

Sequence 2, Appli
Sequence 28765,
Sequence 1160, A

ALIGNMENTS

(without alignments)
10543.293 Million cell updates/sec

RESULT 1
US-09-896-522-2
; Sequence 2, Application US/09896522
; Patent No. US20020055161A1

APPLICANT: Glucksmann, Maria A.

; TITLE OF INVENTION: 57658, A NOVEL HUMAN URIDINE KINASE AND
; TITLE OF INVENTION: USES THEREOF

```

; FILE REFERENCE: 381552001700
; CURRENT APPLICATION NUMBER: US/09/896,522
; CURRENT FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: 60/216,503

```

; NUMBER OF SEQ ID NOS: 6

; SEQ ID NO 2

TYPE: PRT

US-09-896-522-2

Pred. No.:

Percent Similar

```

US-09-896-522-3 (1-834) x US-09-896-522-2 (1-277)

Oy      1  ATGGCTTCGGCGGAGCCGAAGACTGCGAGACCCTGGCGGAGCCGACCGTCCGAC  60
          |||||
Db      1  MetAlaserIaClyGlyGlnApcYcgsGlnSerProAlaProGlnAlaAspArgHis  20

Best Local Similarity: 100.00%  Mismatches: 0
Query Match: 100.00%  Indels: 0
DB: 9  Gaps: 0

```

12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 1041 1042 1043 1044 1045

Db	2.	GlnArgProPheLeuIleGlyValSerGlyThrIleSerGlyLysSerThrValCys	40
Qy	121	GAAAGATCATGGACGTTCTGGGACAGAAACGAGGTGAAACGGCGACGGGAAGGTGTC	180
Db	41	GlnLysIleLeuGluLeuLeuGluGlnGlnGlnValGlnGlnArgGlnArgLysValVal	60
Qy	181	ATCCTGAGCCAGACAGGTTCTTACAAGGTCCTGACGCGACAGACGAAGGCCAAGCCCTTG	240
Db	61	IleLeuSerGlnAspArgPheTyrLysValLeuThrIleGlnGlnLysAlaLysAlaIleu	80
Qy	241	AAGGACAGTACAACTTTGGACATCCAGATGCTTGGTAATGATTTGATGACAGAGCT	300
Db	81	LysGluGlnTyrAsnAspHisPheProAspAlaPheAspHisAspLeuMetHisArgThr	100
Qy	301	CTGAGAACATCGTGGAGGGCAAAACGGGTGAGAGGTGCCAGCTATGATTTTGTGACAC	360
Db	101	LeuLysAsnIleValGlnGluGlySerThrValGlnValProThrTyrAspPheValThrHis	120

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	length	DB	ID	Description
---------------	-------	----------------	--------	----	----	-------------

Db 161 Aspergillus fumigatus ArgValLeuArgValArgGlyArgAspLeu 180
Qy 541 GAGCAGATTCTGACGCACTACCACTTCTGTAAGCCGGCTTCGAGAGTTCTGCTG 600
Db 181 Glutamine leucine threonine threonine valine proline arginine leucine 200
Qy 601 CCGCAAGAAGTATGCGCATGTGATCATCCACAGAGAGGACATATATGTTGGCATC 660
Db 201 Prothymosin ValAlaAspValIleIleProArgValAspAspMetValAlaIle 220
Qy 661 AACCTGATCTGCACACATCTGACATCTGATGATGATGATGATGATGATGATGATG 720
Db 221 AsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysIleTyrHisArg 240
Qy 721 GGAGGCTTCATGGCGGAGCTACACAGGACCTTTTGTAGCCAGGGAGACCCCTGGG 780
Db 241 GlyIleSerAsnGlyArgSerTyrIleArgThrPheSerGluProGlyAspHisProGly 260
Qy 781 ATGCTGACCTCTGGCAACGGTCACTTTGGAGTCCAGCAGACCCAC 831
Db 261 MetLeuThrSerGlyArgSerHisLeuGluSerSerArgProHis 277
RESULT 2
US-10-029-386-28765
Sequence 28765, Application US/10029386
Publication No.: US20030194704A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
FILE REFERENCE: ABOICA-X-2
CURRENT FILING DATE: 2001-12-20
NUMBER OF SEQ ID NOS: 34288
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 28765
LENGTH: 60
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO CHR9.1
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.5
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.8
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.6
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.7
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.5
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.3
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.3
OTHER INFORMATION: SWISSPROT HIT: P52623, EVALU8.00e-26
US-10-029-386-28765
Alignment Scores:
Pred. No.: 6.02e-47 Length: 60
Score: 60.00 Matches: 60
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 21.66% Indels: 0
DB: 12 Gaps: 0
US-09-896-522-3 (1-834) x US-10-029-386-28765 (1-60)
Qy 652 GTTGCATCAACCTGATCTGCAGACATCCAGACATTTGAAATGTGACATTCGAAA 711
Db 1 ValAlaIleAsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysIle 20
Qy 712 TGCGACCGAGAGGCTCCAAATGGGGGAGCTACAAAGCGGACCTTTTTCGACCGAGGAC 771
Db 21 TtrpHisArgGlyArgSerAsnGlyArgSerTyrIleArgThrPheSerGluProGlyAsp 40
Qy 772 CACCTGGAGTGTGACCTCTGGCAACGGTCACTTTGAGTCCAGCAGACGCCAC 831

Db 41 HisProGlyMetLeuThrSerGlyArgSerHisLeuGluSerSerArgProHis 60
RESULT 3
US-09-925-300-1160
Sequence 1160, Application US/09925300
Patent No. US20020151681A1
GENERAL INFORMATION:
APPLICANT: Craig Rosen,
APPLICANT: Steve Ruden,
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
FILE REFERENCE: PA101
CURRENT APPLICATION NUMBER: US/09/925,300
CURRENT FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: PCT/US00/05988
PRIOR FILING DATE: 2000-03-08
PRIOR APPLICATION NUMBER: 60/124,270
PRIOR FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 1890
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1160
LENGTH: 337
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (38)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (46)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (155)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (169)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-300-1160
Alignment Scores:
Pred. No.: 2.05e-14 Length: 337
Score: 25.00 Matches: 25
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 9.03% Indels: 0
DB: 10 Gaps: 0
US-09-896-522-3 (1-834) x US-09-925-300-1160 (1-337)
Qy 565 ACCTGTGTAAGCGGCTTCGAGAGATTCTGCTGCCGACAAAGAATATGCGGATGTG 624
Db 263 ThrPheValIleProAlaPheGlnGluPheCysLeuProThrIleValAlaAspVal 282
Qy 625 ATCATCCACGAGAGA 639
Db 283 IleIleProArgGly 287
RESULT 4
US-09-896-522-5
Sequence 5, Application US/09896522
Patent No. US20020055161A1
GENERAL INFORMATION:
APPLICANT: Glucksmann, Maria A.
TITLE OF INVENTION: 57658, A NOVEL HUMAN URIDINE KINASE AND
FILE REFERENCE: 381552001700
CURRENT APPLICATION NUMBER: US/09/896,522
CURRENT FILING DATE: 2001-06-28
PRIOR APPLICATION NUMBER: 60/216,503
PRIOR FILING DATE: 2000-06-30
NUMBER OF SEQ ID NOS: 6
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 5

```

; LENGTH: 125
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Consensus amino acid sequence
US-09-896-522-5

```

Alignment Scores:

Pred. No.:	4,83e-05	Length:	125
Score:	15.00	Matches:	15
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	5.42%	Indels:	0
DB:	9	Gaps:	0

US-09-896-522-3 (1-834) x US-09-896-522-5 (1-125)

```

QY      736 CGAGCTACAAGCGGACCTTTCTGAGCCAGGAGGACCCCTGAG 780
      |||||
Db      94 ArgSerTrpLysArgThrPheSerGluProGlyAspHisProGly 108

```

Search completed: November 25, 2003, 08:04:08
 Job time : 30.1798 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: November 25, 2003, 07:43:14 ; Search time 15.7775 Seconds
(without alignments)
10166.994 Million cell updates/sec

Title: US-09-896-522-3
Perfect score: 277
Sequence: 1 atgagctcgcgagcgagcgca.....ccagcagcagaccacactga 834

Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 283308 seqs, 96168682 residues

Word size: 15

Total number of hits satisfying chosen parameters: 0

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: listing first 45 summaries

Command line parameters:

-MODE=frame+n2p.model -DEV=xlh
-Q=/cgr2_1/USPTO_spool/US0896522/runatc_21112003_184145_3234/app_query.faeta_1.2830
-DB=PIR_76 -QFMT=faetan -SUFFIX=n2p_ol1.rpr -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=quality -THR MIN=15 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc
-NORM=ext -HEAPSIZE=500 -MAXLEN=200000000
-USER=US0896522.ccgcn_1_1_68 @runatc_21112003_184145_3234 -NCPU=6 -ICPU=3
-NO_MMAPP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : PIR_76.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Query Length	ID	Description
No matches found				

Search completed: November 25, 2003, 07:59:52
Job time : 15.7775 secs

This Page Blank (uspto)

XX	The invention relates to human nucleic acids (AA15798-AA161369) and
CC	the encoded polypeptides (AAM39842-AAM42213) with neotropic,
CC	immunosuppressant and cytostatic activity. The polynucleotides are useful
CC	in gene therapy. A composition containing a polypeptide or polynucleotide
CC	of the invention may be used to treat diseases of the peripheral nervous
CC	system, such as peripheral nervous injuries, peripheral neuropathy and
CC	localized neuropathies and central nervous system diseases, such as
CC	Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC	lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC	utilization of the activities such as: Immune system suppression,
CC	Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC	and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC	assays for receptor activity, arthritis and inflammation, leukaemia and
CC	C.N.S disorders.
CC	Note: The sequence data for this patent did not form part of the printed
CC	specification.
CC	
XX	
SQ	Sequence 277 AA;
 Alignment Scores:	
Pred. No.:	5,11e-272
Score:	277.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	100.00%
DB:	22
	Gaps: 0
US-09-896-522-3 (1-834) x AAM39502 (1-277)	
OY	1 ATGGCTTCGGCGGAGGCGGAACACTGCAGAGCCCCGGCGCGGACCGGTCCGAC 60
Db	1 MeelalaserlaaglyglnumprCyvsgIuseerProalaProglnalaAspArgprohis 20
OY	61 CAGCGGCCCTTCTGTATAGGGGGTGAGCGCGGCACTGCCAGCGGGAAATGCACCSTGTNGT 120
Db	21 GlnhrgProheuleulleglyValsercIyglThrlasergIlylssetThlValCys 40
OY	121 GAAGAATCATGAGATTGCTGGGACAAGAACAGGTGGAAACAGCGGCAGCGGAAGTGGTC 180
Db	41 GlulysllewetsIleueugIleugIlnaInglnuValGlugInargGlnatrgLvalVal 60
OY	181 ATCTGAGCCAGGAKOAGGTTTTCACAGAGTCTCGAGGCGAGAGCAAGACCAAGGSCCTTG 240
Db	61 IleusenertImaprtgPheutytllysValLeuthrilaaglInlyrsalalyalaieu 80
OY	241 AAAAGCACGTACAAATTTTTCACATCCAGATGCCCTTGATATGATTGATGCACAGACT 300
Db	81 LyeGIylntIrtaenphearsphIsprodarllaRheaAspaenArlemehnaIgThr 100
OY	301 CTGAAGAAKATCGTGGAGGGCAAACAGGTGGAGGTGCGSASTATGATTTTGTGACAC 360
Db	101 LeulyasnnlleValGlugIlystrValaGlualProthrTyAspPheValThnIs 120
OY	361 TCAAGGTTTACAGAGACCAAGGAGTCACTCCGCGAGCGTGGTCTGTTTGAGGGCATC 420
Db	121 SerhrgIeurProglutlrThrValaValItyrrolaAsparValaLeuheuhgluyile 140
OY	421 TTGGTGTTCACAGCCAGAGATCCGGACATGTTCCATGTGCGCTTCTGTGGACACC 480
Db	141 LeuValPheTySerGingluilleatrgAspmetPrenhIsLeuatrgLeuPheValAspThr 160
OY	481 GAGTCGAGCTCAGGGGTCTGGAAGATTCCTCGGACAGTGGCGCGCGGAGGAGACTG 540
Db	161 AspsertAspAlarIrgueusertArgavaIlseuIdgAspValAdgaVgIlyAspLeu 180
OY	541 GAGCAGATTTCACGAGTACACACACCTTCGTAAGCGCGGCTTTCAGAGATTTCGCTG 600
Db	181 GlugInlleuthrghIntyrThrThrPheValaValysprollaRheglInuRheCyleu 200
OY	601 CCAGCAAAAAGATATGCGCATGTGATCATCCACAGAGAGTGGACATAATGTGGCATC 660
Db	201 ProthrIysIysIyrlaAspValilleleProldvgIlyValaAspAsmeValaAlle 220

QY	666	AACTGATGGTGGAGACATCCAGGACATTTCGAAATGGTGCACATCTGCAATGGACACGA	72
Db	221	AsnLeuIleValGlnHisIleGlnMetPilleuSngIyaPrlIcYbLysTrpHisArg	240
QY	721	GGAGGGGTCCAAATGGGCGGAGCTTACCAAGGGGACCTTTTCTGAGCCAGGGGACCAACCTGGG	780
QY	781	ATGCTGACCTCTGGGCAACGCTCACATTTGGAGTCCAGACAGACCCAC	831
Db	261	MetLeuThrSerGlyLysArgSerHisIleuGlnSerSerSerArgProHis	277
RESULT 2			
AAAB93941	ID	AAAB93941 standard; Protein; 277 AA.	
AAAB93941;			
AC	26-JUN-2001	(first entry)	
DT			
XX			
DE		Human protein sequence SEQ ID NO:13952.	
XX			
KW		Human; primer; detection; diagnosis; antisense therapy; gene therapy.	
XX			
OS		Homo sapiens.	
XX			
PN		EPI074617-A2.	
PD		07-FEB-2001.	
XX			
PF		28-JUL-2000; 2000EP-0116126.	
XX			
PR		29-JUL-1999; 99JP-0248036.	
PR		27-AUG-1999; 99UP-0300253.	
PR		11-JAN-2000; 2000JP-0118776.	
PR		02-MAY-2000; 2000JP-0183767.	
PR		09-JUN-2000; 2000JP-0241899.	
XX			
PA		(HELI-) HELIX RES INST.	
XX			
P1		Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;	
P1		Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;	
XX			
XX		WPI; 2001-318749/34.	
PT		Primer sets for synthesizing polynucleotides, particularly the 5602	
PT		full-length cDNAs defined in the specification, and for the detection	
PT		and/or diagnosis of the abnormality of the proteins encoded by the	
PT		full-length cDNAs -	
XX			
BS		Claim 8; SEQ ID 13952; 2537bp + CD ROM; English.	
XX			
CC		The present invention describes primer sets for synthesizing 5602	
CC		full-length cDNAs defined in the specification. Where a primer set	
CC		comprises: (a) an oligo-dT primer and an oligonucleotide complementary	
CC		to the complementary strand of a polynucleotide which comprises one of	
CC		the 5602 nucleotide sequences defined in the specification, where the	
CC		oligonucleotide comprises at least 15 nucleotides; or (b) a combination	
CC		of an oligonucleotide comprising a sequence complementary to the	
CC		complementary strand of a polynucleotide which comprises a 5'-end	
CC		sequence and an oligonucleotide comprising a sequence complementary to a	
CC		polynucleotide which comprises a 3'-end sequence, where the	
CC		oligonucleotide comprises at least 15 nucleotides and the combination of	
CC		the 5'-end sequence/3'-end sequence is selected from those defined in	
CC		the specification. The primer sets can be used in antisense therapy and	
CC		in gene therapy. The primers are useful for synthesizing polynucleotides,	
CC		particularly full-length cDNAs. The primers are also useful for the	
CC		detection and/or diagnosis of the abnormality of the proteins encoded by	
CC		the full-length cDNAs. The primers allow obtaining of the full-length	
CC		cDNAs easily without any specialised methods. AAH03166 to AAH13628 and	
CC		AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to	
CC		AAH95893 represent human amino acid sequences; and AAH15629 to AAH15632	

CC represent oligonucleotides, all of which are used in the exemplification
 CC of the present invention.

XX Sequence 277 AA;

Alignment Scores:

Score: 5,11e-272 Length: 277
 Pred. No.: 277.00 Matches: 277
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 22 Gaps: 0

US-09-896-522-3 (1-834) x AAB93941 (1-277)

QY 1 ATGGCTTGGCGGAGGCGAAGACTGCGAGACCCCGCGGAGCCGACCGTCCGAC 60
 Db 1 MetAlaSerIleAlaGlyGluAspCysGluSerProAlaProGluAlaAspArgProHis 20
 QY 61 CAGCGGCCCTTCCGATAGGGGTGAGCGGGGCACTGCCGCGGAAAGTGCAGCTGTGT 120
 Db 21 GlnArgProPheLeuIleGlyValSerGlyThrAlaSerGlyLysSerThrValCys 40
 QY 121 GAGAGATCATGAGATTGCTGGGACAGAACGAGGTGGACAGCGCGCAGCGAAAGTGTGTC 180
 Db 41 GluLysIleMetGluLeuLeuGlyGlnAsnGluValGluGlnArgGlnArgLysValVal 60
 QY 181 ATCTGAGCCAGAGACAGGTCTTACAGAGTCTTGCAGCGACAGCAAGAGCCGCTTG 240
 Db 61 IleLeuSerGlnAspArgPheTyrLysValLeuThrIleGluGlnLysAlaLysAlaLeu 80
 QY 241 AAAGSACAGTACATTTGACCATCCAGATCCGCTTGATATGATTTGATGCAGGACT 300
 Db 81 LysGlyGlnIlyrAsnPheAspHisProAspAlaPheAspAsnLeuMetHisArgThr 100
 QY 301 CTGAGAACATCGTGGAGGCGCAAAACGCTGAGAGGTGCGCAGCTATGATTTTGACACAC 360
 Db 101 LeuLysAsnIleValGluGlyLysThrValGluValProThrLysAspPheValThrHis 120
 QY 361 TCAAGTTACCAAGAACACACGCTGCTTACCTCGCGGACGCTGTTCTGTTGAGGCGATC 420
 Db 121 SerArgLeuProGluThrThrValIleTyrProAlaAspValIleuPheGluGlyIle 140
 QY 421 TTGTTGTTACACCGCAGGAGATCCGCGGACATGTTCCACTGGCCCTTCGTCGACACC 480
 Db 141 LeuValPheTyrSerGlnGluIleArgAspMetPheHisLeuArgLeuPheValAspThr 160
 QY 481 GACTCCGACGTCAGGCTGTCTCGAAGAGTCTCCGGGACGTCGCGGAGGAGGAGACTG 540
 Db 161 AspSerAspValArgLeuSerArgValIleuArgAspValArgArgGlyAlaGlyAspLeu 180
 QY 541 GAGCAGATTCTGACGACGATACACCACTTGTGAAGCGGCTTCGAGAGTTCTGCTG 600
 Db 181 GluGlnIleLeuThrGlnIlyrThrThrPheValLysProAlaPheGluGluPheCysLeu 200
 QY 601 CGCAGAAAGAGTATGCCGATGTGATCATCCGACGAGAAATGACATATAGTTGTCATC 660
 Db 201 ProThrLysLysThrAlaAspValIleIleProArgGlyValAspAsnMetValAlaIle 220
 QY 661 AACCTGATCGTGACGACATCCAGGACATTTGTAATGATGACATCTGCAATGGACCGA 720
 Db 221 AsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyLysPheCysIlyrThrPheArg 240
 QY 721 GAGAGGTCATAGGCGGAGCTACAGCGGACCTTTCTGAGCGAGGAGCAACCTCGGG 780
 Db 241 GlyGlySerAsnGlyArgSerTyrLysArgThrPheSerIleuProGlyAspHisProGly 260
 QY 781 ATGCTGACCTCTGGCAAGCGTCACTTTGAGATCCAGGACGAGCCCGAC 831
 Db 261 MetLeuThrSerGlyLysArgSerHisLeuGluSerSerArgProHis 277

ID AAE16592 standard; Protein; 277 AA.
 XX AAE16592;
 AC 18-APR-2002 (first entry)
 XX Human 57658 protein.
 DT Human 57658 protein.
 XX
 XX
 DE Human; uridine kinase-like protein; hematopoietic neoplastic disorder;
 XX 57658 protein; leukemia; hematopoietic disorder; diabetes mellitus;
 KW rheumatoid arthritis; psoriasis; dermatitis; osteoporosis; rickets;
 KW sarcoma; myocardial infarction; hypertension; atherosclerosis;
 KW Alzheimer's disease; Parkinson's disease; Huntington's disease;
 KW gene therapy; chromosome mapping; tissue typing; dermatological;
 KW cytoskeletal; osteopathic; cardiac; neuroprotective; nootropic;
 KW anticovulant.
 XX
 XX Homo sapiens.
 OS
 XX
 XX
 FH Key
 FT Modified-site
 FT Location/Qualifiers
 FT 5..10
 FT /label= N-myristoylation_site
 FT 25..151
 FT /note= "Kinase uridine transferase ATP-binding
 FT phosphoribulokinase monophosphokinase precursor PRK
 FT cycle phosphopentokinase domain"
 FT 25..231
 FT /note= "Phosphoribulokinase domain; Protein kinase
 FT family domain"
 FT 27..32
 FT /label= N-myristoylation_site
 FT 30..37
 FT /label= ATP/GTP-binding_site
 FT 34..36
 FT /label= Protein_kinase-C_phosphorylation_site
 FT 38..41
 FT /label= Casein_kinase_II_phosphorylation_site
 FT 100..102
 FT /label= Protein_kinase-C_phosphorylation_site
 FT 108..115
 FT /label= Tyrosine_kinase_phosphorylation_site
 FT 154..277
 FT /note= "Kinase uridine monophosphokinase transferase
 FT ATP-binding kinase-like ribonucleoside pyrimidine F1S
 FT cDNA domain"
 FT 160..163
 FT /label= Casein_kinase_II_phosphorylation_site
 FT 167..169
 FT /label= Protein_kinase-C_phosphorylation_site
 FT 202..204
 FT /label= Protein_kinase-C_phosphorylation_site
 FT 241..246
 FT /label= N-myristoylation_site
 FT 247..249
 FT /label= Protein_kinase-C_phosphorylation_site
 FT 251..254
 FT /label= Casein_kinase_II_phosphorylation_site
 FT 260..265
 FT /label= N-myristoylation_site
 FT 264..266
 FT /label= Protein_kinase-C_phosphorylation_site
 FT 264..267
 FT /label= Amidation_site
 FT 268..271
 FT /label= Casein_kinase_II_phosphorylation_site
 FT 273..275
 FT /label= Protein_kinase-C_phosphorylation_site
 XX
 XX
 PN MO200202761-A2.
 XX 10-JAN-2002.
 PD 28-JUN-2001; 2001MO-US21063.
 XX

XX 30-JUN-2000; 2000US-216503P.
 PR (MILL-) MILLENNIUM PHARM INC.
 XX Gluckemann MA;
 PI WPI: 2002-140091/16.
 XX N-PSDB; AAD27186.
 DR
 XX
 XX New isolated human uridine kinase family polypeptide 57658, useful for
 PT treating hematopoietic neoplastic disorders and disorders of neurons,
 PT heart and blood vessels -
 XX
 XX Claim 4; Fig 1a; 103pp; English.
 PS
 XX The patent discloses human uridine kinase-like polypeptides, designated
 CC 57658 and polynucleotides encoding such proteins. 57658 DNAs and proteins
 CC are useful for developing diagnostic and therapeutic agents for 57658-
 CC mediated or related disorders such as hematopoietic neoplastic disorders
 CC (e.g. leukemia), hematopoietic disorders (e.g. psoriasis, dermatitis,
 CC diabetes mellitus, rheumatoid arthritis, disorders of bone (e.g. rickets,
 CC osteoporosis), heart (e.g. sarcoma, myocardial infarction), blood vessels
 CC (e.g. hypertension, atherosclerosis) and neurons (e.g. Huntington's
 CC disease, Alzheimer's disease, Parkinson's disease). 57658 sequences or
 CC their antibodies are useful in screening assays, detection assays (e.g.
 CC forensic biology), and predictive medicine (e.g. diagnostic assays,
 CC prognostic assays, and monitoring clinical trials and pharmacogenomics).
 CC They are useful as reagents for diagnosing and treating 57658-mediated
 CC disorders. 57658 DNAs are useful in gene therapy, in chromosome mapping,
 CC to identify an individual from a minute biological sample (tissue typing)
 CC and to aid in forensic identification of the biological sample. The
 CC present sequence is human 57658 protein.
 CC
 XX
 SQ Sequence 277 AA;
 Alignment Scores:
 Pred. No.: 5,11e-272 Length: 277
 Score: 277.00 Matches: 277
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: Gaps: 0
 US-09-896-522-3 (1-834) x AAE16592 (1-277)
 QY 1 ATGGCTTCGGGCGGAGGCGAAGACTGCGAGAGCCCGCGCGAGCGGACCGCTCGCAC 60
 DB 1 MetAlaSerAlaGlyGlyGluAspCysGluSerProAlaProGluAlaAspArgProHis 20
 QY 61 CAGGGGCGCTTCCTGATGAGGGGTGAGCGGGCGCACTGCGACGGGAAGTGCACCGTGTGT 120
 DB 21 GlnArgProPheLeuIleGlyValSerGlyGlyThrAlaSerGlyLysSerThrValCys 40
 QY 121 GAGAGATCATGAGTGTCTGGACAGAACAGAGGTGGAACAGCGGCGGAGGAGTGTGTC 180
 DB 41 GluValIleMetGluLeuLeuGlyGlnAsnGluValGluGlnArgValValVal 60
 QY 181 ATCTGAGCCAGAGCAGGTTCTACAGAGTCTGACGCGACAGAGCAAGGCCAAGCCTTG 240
 DB 61 IleLeuSerGlnAspArgPheTyrLysValLeuThrAlaGluGlnLysAlaLysAlaLeu 80
 QY 241 AAAGAGACATGATCTTTTGGACATCCAGATGCTTGTGATATGATTTGATGACAGACT 300
 DB 81 LysGlyGlnTyrAsnPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThr 100
 QY 301 CTGAAGACATCGTGGAGGCGAAGAGCGTGGAGGTGCGACCTTGTATTTTGTGACACAC 360
 DB 101 LeuLysAsnIleValGluGlyLysThrValGluValProThrTyrAspPheValThrHis 120
 QY 361 TCAAGGTTACAGAGACCGAGTGTCTACCTCGGACGTCGTCTCTGTTTGAAGGACATC 420
 DB 121 SerArgLeuProGluTThrThrValValTyrProAlaAspValValLeuPheGluGlyIle 140

QY 421 TTGGTGTCTACAGCCAGAGAGATCCGGGACATGTTCCACTGCGCCTTTCGTGACACC 480
 DB 141 LeuValPheTyrSerGlnGluIleArgAspMetPheHisLeuArgLeuPheValAlaPthr 160
 QY 481 GACTCCAGAGTCAGAGCTGTCTCGAAGAGTTCCTCCGGAAGTGGCGGAGAGGAGACTG 540
 DB 161 AspSerAspValArgLeuSerArgValIleuArgAspValArgArgGlyArgAspLeu 180
 QY 541 GACCAATTTCTGACGCGATACACACACTTCGTGAAGCCGGCCTTCGAGAGATTCTGCTG 600
 DB 181 GluGlnIleLeuThrGlnTyrThrThrPheValLysProAlaPheGluIleuPheCysLeu 200
 QY 601 CCGAGAAAGAGATGCCGATGATCATCCACGAGAGAGTGGAGCAATATGATGTCATC 660
 DB 201 ProThrLysLysTyrAlaAspValIleIleProArgGlyValAspAsnMetValAlaIle 220
 QY 661 AACCTGATCGTGCAGACATCCAGACATTTGTAATGTGACATCTGCAATGGACCGA 720
 DB 221 AsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysLysTrpHisArg 240
 QY 721 GAGGGTCCATGGGCGGAGCTACAAAGCGACCTTTCTGAGCCAGGAGGACACCTTGG 780
 DB 241 GlyGlySerAsnGlyArgSerTyrLysArgThrPheSerGluProGlyLysPheHisProGly 260
 QY 781 ATGCTGACCTTCGGCAAGAGTCACTTTGGAGTCCAGAGACACCCGAC 831
 DB 261 MetLeuThrSerIleLysArgSerHisLeuGluSerSerArgProHis 277
 RESULT 4
 ID AAM41288
 AC AAM41288; standard; protein; 296 AA.
 XX
 XX AAM41288;
 AC
 XX
 XX 22-OCT-2001 (first entry)
 DT
 XX
 XX Human polypeptide SEQ ID NO 6219.
 DE
 XX
 XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemoclastic;
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 KW leukemia.
 KW
 XX Homo sapiens.
 OS
 XX
 XX MO20015312-A1.
 PN
 XX
 XX 26-JUL-2001.
 PD
 XX
 XX 26-DEC-2000; 2000MO-US34263.
 PF
 XX
 XX 21-JAN-2000; 2000US-0488725.
 PR 25-APR-2000; 2000US-0552317.
 PR 09-JUL-2000; 2000US-0598042.
 PR 19-JUL-2000; 2000US-0620312.
 PR 03-AUG-2000; 2000US-0653450.
 PR 14-SEP-2000; 2000US-0662191.
 PR 19-OCT-2000; 2000US-0693036.
 PR 29-NOV-2000; 2000US-0727344.
 XX
 XX (HYSE-) HYSEQ INC.
 PA
 XX
 XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Wehman T, Xu C, Xue AJ, Yang Y, Zhang J;
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
 XX
 XX WPI: 2001-442253/47.
 DR N-PSDB; AAI60444.
 XX
 XX Novel nucleic acids and polypeptides, useful for treating disorders

PT such as central nervous system injuries -
 XX
 PS Example 2; SEQ ID NO 6219; 10078bp; English.
 XX
 CC The invention relates to human nucleic acids (AA157798-AA161369) and
 CC the encoded polypeptides (AA438642-AA442213) with nootropic,
 CC immunosuppressant and cytoskeletal activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemias and
 CC C.N.S disorders.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification.
 CC
 XX
 SQ Sequence 296 AA;

Alignment Scores:
 Pred. No.: 5.07e-272 Length: 296
 Score: 277.00 Matches: 277
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 22 Gaps: 0

US-09-896-522-3 (1-834) x AAM41288 (1-296)

QY 1 ATGGCTTCGGCGGAGGCGAAGACTGCGAGAGCCCGCGGAGCGGACCGTCCGAC 60
 Db 20 MetAlaSerAlaGlyGlyGluAerCySgluSerProAlaProGluAlaAerProHis 39
 QY 61 GACGGCGCCCTTCGATGAGGGGTGAGGGGCGACGCGCGGAGGAGTGCCTGCT 120
 Db 40 GluArgProGluLeuIleGlyValSerGlyGlyThrAlaSerGlyLysSerThrValCys 59
 QY 121 GAGAAGATCATGAGTGTCTGGGACAGAAAGAGGTGAAAGCGGCGGAGAGTGTGTC 180
 Db 60 GluLysIleMetGluLeuLeuGlyGlnAsnGluValGluGlnArgGlnAlaGlyValVal 79
 QY 181 ATCTTGAGCGAGGACAGGCTTCTACAGGTCCTGACGGCAGAGCAAGAGCGCTTG 240
 Db 80 IleLeuSerGlnAspArgPheTyrLysValIleuThrAlaGluGlnLysAlaLeu 99
 QY 241 AAGGACAGTACAAATTTTGCATCCAGATGCTTGTATATGATTTGATGACAGACT 300
 Db 100 LysGlyGlnTyrAsnPheAspHisProAspAlaPheAspAsnSpleuMetHisArgThr 119
 QY 301 CTGAAGACATCGTGGAGGGCAAAACGGTGGAGGTGGCGACTATGATTTTGGACAC 360
 Db 120 LeuLysAsnIleValGluGlySerThrValGluValProThrTyrAspPheValThrHis 139
 QY 361 TCAAGTTACAGAGACACAGCGGTGTCTACCTTCGCGACGCTGTCTTGTGAGGCATC 420
 Db 140 SerArgLeuProGluThrValThrValValTyrProAlaAspValIleuPheGluGlyIle 159
 QY 421 TTGGTGTTCACAGCCGAGGATCCGGGACATGTTCCACTGGCGCTTCTGTGACACC 480
 Db 160 LeuValPheTyrSerGlnGluIleArgAspMetPheIleuArgLeuPheValAspThr 179
 QY 481 GACTCCGACGTCAAGCTGTCTCGAAGAGTCTCCGGGACGTGGCGGAGGAGGACCTG 540
 Db 180 AspSerAspValArgLeuSerArgArgValIleuArgAspValArgArgGlyArgAspLeu 199
 QY 541 GAGCAGATTCTGACGACGATACACCACTTGTGAAAGCGGCTTTCGAGAGATTCTGCTG 600
 Db 200 GluGlnIleuThrGlnTyrThrThrPheValLysProAlaPheGluGluPheCysLeu 219

QY 601 CCGAAGAAAGATGATCCGATGATCATCCACGAGAGTGAAGATGATGTTGCCATC 660
 Db 220 ProThrLysLysTyrAlaAspValIleIleProArgGlyValAlaAspAsnMetValAlaIle 239
 QY 661 AACCTGATCGTGCAGACATCCAGACATTTGAAATGTGACATCTGCAATGGACCGA 720
 Db 240 AsnLeuIleValGlnHisIleGlnIleAspIleLeuAsnGlyAspIleCysLysTrpHisArg 259
 QY 721 GAGGGTCCAAATGAGGGGAGCTACAAAGCGGACCTTTTTCAGGCGAGGAGCCCTGGG 780
 Db 260 GlyGlySerAsnGlyArgSerTyrLysArgThrPheSerGluProGlyAspHisProGly 279
 QY 781 ATGCTGACCTTCGGCAAGCGTACATTTGAGTCCAGACGAGACCCAC 831
 Db 280 MetLeuThrSerGlyLysArgSerHisLeuGluSerSerArgProHis 296
 RESULT 5
 ID AA014412
 AC AA014412
 DE 02-MAY-2002 (first entry)
 XX Protein of a human uridine kinase (UDK).
 DE
 XX
 KW Human; uridine kinase; diagnostic assay; mutation detection; UDK;
 KW probe; chromosome localisation study; tissue expression; gene therapy;
 KW antibody; vaccine; human ovarian cancer; immunological disorder;
 KW human colon carcinoma; immunogen.
 XX
 OS Homo sapiens.
 XX
 PN WO200172963-A2.
 XX
 PD 04-OCT-2001.
 XX
 PF 27-MAR-2001; 2001MO-US09663.
 XX
 PR 27-MAR-2000; 2000US-0536647.
 XX
 PA (SMIK) SMITHKLINE BEECHAM CORP.
 XX
 PI Ho YS, Johnson RK;
 XX
 DR WPI; 2001-626259/72.
 DR N-Psdb; AAK98735.
 XX
 PT Novel human uridine kinase polypeptides useful for treating cancers,
 PT and to identify agonists and antagonists of the polypeptide useful for
 PT treating conditions associated with uridine kinase imbalance -
 XX
 PS Claim 3; Page 29-30; 31pp; English.
 XX
 CC The invention relates to newly identified human uridine kinase (UDK)
 CC polypeptides and polynucleotides and methods for producing such
 CC polypeptides by recombinant techniques. Also disclosed in the invention
 CC are methods for utilising uridine kinase polypeptides and polynucleotides
 CC in diagnostic assays. The polynucleotides and polypeptides of the
 CC invention may be used as diagnostic reagents by detecting mutations in an
 CC associated gene. An array of oligonucleotide probes comprising the
 CC uridine kinase polynucleotide sequence or fragments thereof can be
 CC constructed to conduct efficient screening of genetic mutations, for
 CC example. Detection of abnormally decreased or increased levels of
 CC polypeptide or mRNA expression may also be used for diagnosing or
 CC determining susceptibility of a subject to a disease of the invention.
 CC The polynucleotide sequences of the invention can be used for chromosome
 CC localisation studies and tissue expression studies. The polypeptides of
 CC the invention or fragments thereof may be used as immunogens to produce
 CC antibodies. These antibodies may be employed to isolate or identify
 CC clones expressing the polypeptide. The polypeptides and polynucleotides
 CC of the invention can be used as a vaccine or in gene therapy to treat
 CC diseases such as human ovarian cancer, human colon carcinomas, and

CC immunological disorders. This sequence represents the protein of a human
 CC uridine kinase of the invention.

XX Sequence 260 AA;

Alignment Scores:

Pred. No.: 9,2e-255 Length: 260
 Score: 260.00 Matches: 260
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 93.86% Indels: 0
 DB: 22 Gaps: 0

US-09-896-522-3 (1-834) x AA014412 (1-260)

QY 1 ATGGCTTCGGGGGAGGAGGAGATCGGAGAGCCCGGCGGAGCCGACCTGCGCAC 60
 Db 1 MetAlaSerAlaGlyGlyGluAspCysGluSerProAlaProGluAlaAspArgProHis 20
 QY 61 CAGGGCCCTTCCTGATAGGGGTGAGCGGCGGCACTGGCAGCGGGAAGTGCAGCGTGAT 120
 Db 21 GlnArgProPheLeuIleGlyValSerGlyThrAlaSerGlyLysSerThrValCys 40
 QY 121 GAGAGATCATGAGTGTCTGGACAGAACGAGGTGAAACGCGGACGCGAGCGAGTGTGTC 180
 Db 41 GluLysIleMetGluLeuLeuGlyGlnAsnGluValGluGlnArgGlnArgLysValVal 60
 QY 181 ATCTGGAGCCGAGCAGGTTCTACAGGCTCTGACGCGAGAGCGAAGGCGCAAGCGCTTG 240
 Db 61 IleLeuSerGlnAspArgPheTyrlsValValLeuThrAlaGluGlnLysAlaLysAlaLeu 80
 QY 241 AAAGACAGTCAATTTTGAACATCCAGATGCTTTGATTAATGATTGATCAGAGACT 300
 Db 81 LysGlyGlnIlePheAsnAspHisProAspAlaPheAspAlaPheLeuMetHisArgThr 100
 QY 301 CTGAAGAACATCGTGGAGGGGAAAACGGTGGAGGTGCGGACCTATGATTGTCACACAC 360
 Db 101 LeuLysAsnIleValGluGlyLysThrValGluValProThrTyraAspPheValThrHis 120
 QY 361 TCAGAGTTACCGAGACACGCGTGTCTCAACCTCGGACGCGTGTCTGTTGAGGCGATC 420
 Db 121 SerArgLeuProGluThrThrValValValTyrlProAlaAspValValLeuPheGluGlyIle 140
 QY 421 TTGGTGTCTCACAGCCAGAGATCCGGGACATGTTCCAGCTGCGCTCTTGTCGACAC 480
 Db 141 LeuValPheTyrlSerGlnGlnIleArgAspMetPheHisLeuAlaGlyLeuPheValAspThr 160
 QY 481 GACTCCGACGTCAGGCTGTCTCGAAGAGTTCTCCGGGACGTGCGCCGAGGAGGAGACTG 540
 Db 161 AspSerAspValAlaArgLeuSerArgValLeuAlaArgValAlaArgGlyArgAspLeu 180
 QY 541 GAGCAGATTCTGACGAGTACACACCTTCGTGAAGCGCGCTTCGAGAGTTTGCTG 600
 Db 181 GluGlnIleLeuThrGlnIleTyrlThrPheValLysProAlaPheGlnGluPheCysLeu 200
 QY 601 CCGACAAAGAGTATGCGCATGTGATCATCCACGAGAGTGACAAATATGTTGCCATC 660
 Db 201 ProThrLysLysTyrlAlaAspValIleIleProArgGlyValaAspAsnMetValAlaIle 220
 QY 661 AACCTGATCGTGACGACATCCAGACACATTCGATGAGTGACATCTGCAATGGCACCGA 720
 Db 221 AsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysLysThrPheArg 240
 QY 721 GGAGGGTCGCAATGGGCGGAGTACAAGCGGACCTTTCTGAGCCAGGGAGCAACCTGGG 780
 Db 241 GlyGlySerAsnGlyArgSerTyrlsArgThrPheSerGluProGlyAspHisProGly 260

RESULT 6
 ID AA64506
 AC AA64506; standard; Protein, 276 AA.

DT 02-OCT-2001 (first entry)
 XX
 DE Human uridine kinase.
 XX
 KW Human; uridine kinase; UK.
 XX
 OS Homo sapiens.
 XX
 PN CN1287172-A.
 XX
 PD 14-MAR-2001.
 XX
 PF 07-SEP-1999; 99CN-0118818.
 XX
 PR 07-SEP-1999; 99CN-0118818.
 XX
 PA (UYFU-) UNIV FUDAN.
 XX
 PI Yu L, Zhao Y, Zhang H;
 XX
 DR WPI: 2001-409529/44.
 DR N-PSDB; AAH75355.
 XX
 PT Human uridine kinase and its coding sequence, preparation and
 PT application -
 XX
 PS Claim 2; Page 15-16 (Disclosure); 20pp; Chinese.
 XX
 CC The invention relates to human uridine kinase (UK).
 XX
 SQ Sequence 276 AA;

Alignment Scores:
 Pred. No.: 1.04e-182 Length: 276
 Score: 189.00 Matches: 189
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 68.23% Indels: 0
 DB: 22 Gaps: 0

US-09-896-522-3 (1-834) x AA64506 (1-276)

QY 172 AAGGTGTCATCTCGAGCCAGACAGAGTCTTCAAGGTCCTGACGCGACAGAGGCC 231
 Db 57 LysValValIleLeuSerGlnAspArgPheTyrlsValValLeuThrAlaGluGlnLysAla 76
 QY 232 AAGGCTTGAAGAGACAGTCAATTTTGAACATCCAGATGCTCTGATTAATGATTGATG 291
 Db 77 LysAlaLeuLysGlyGlnIlePheAsnAspHisProAspAlaPheAspAsnAspLeuMet 96
 QY 292 CACAGACCTGTAAGAACATCGTGGAGGGGCAAAACGGTGGAGGTGCGGACCTATGATTTT 351
 Db 97 HisArgThrLeuLysAsnIleValGluGlyLysThrValGluValProThrTyraAspPhe 116
 QY 352 GTGACACACTCAAGGTTTACAGACACAGAGTGTGTTCACTGCGGACGTGTTCTGTTT 411
 Db 117 ValThrHisSerArgLeuProGluThrThrValValTyrlProAlaAspValValLeuPhe 136
 QY 412 GAGGGCATCTTGTTGTTCTTACAGCCAGAGATCCGGGACATGTTCCACTGCGCTCTTC 471
 Db 137 GluGlyIleLeuValPheTyrlSerGlnGlnIleArgAspMetPheHisLeuAlaGlyLeuPhe 156
 QY 472 GTGACACGAGCTCGAGCGTCAAGGCTGTCTCGAAGAGTTCTCCGGGACGTGCGGCGAGGG 531
 Db 157 ValAspThrAspSerAspValAlaArgLeuSerArgValLeuArgAspValAlaArgGly 176
 QY 532 AGGACCTGGAGCAGATTCTGACGACGATACACCACTTGTGAAGCGCGGCTTGCAGAG 591
 Db 177 ArgAspLeuGluGlnIleLeuThrGlnIleTyrlThrPheValLysProAlaPheGlnGlu 196
 QY 592 TTCTGCTCGCGCAAAAGAGTATGCGCATGTGATCATCCACGAGAGTGACAAATATG 651
 Db 197 PheCysLeuProThrLysLysTyrlAlaAspValIleIleProArgGlyValaAspAsnMet 216

QY 652 GTTCCATCAACCTGATCGTGACACATCCAGAGATTGATGTCATCTGCAAA 711
 |||||
 Db 217 VALAIIeAhsmuilevalaIghIhIleInspIleuAhngIlyAspIleCysLys 236
 |||||
 QY 712 TGGCACCAGAGAGGTCCAATGGCGCG 738
 |||||
 Db 237 TrpHIAArgGLyGLySerAengIlyArg 245
 |||||
 RESULT 7
 ABB89353
 ID ABB89353 standard; Protein; 190 AA.
 XX ABB89353;
 AC ABB89353;
 XX
 DT 24-MAY-2002 (first entry)
 XX
 DE Human polypeptide SEQ ID NO 1729.
 XX
 XX Cytostatic; immunosuppressive; nocotropic; neuroprotective; antiviral;
 KM antileptergic; hepatotropic; antidiabetic; antiinflammatory; antitumor;
 KM vulnerrary; anticonvulsant; antibacterial; antifungal; antiparasitic;
 KM cardiac; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KM neurological disease; infection; human; secreted protein.
 XX
 OS Homo sapiens.
 XX
 PN WO200190304-A2.
 XX
 PD 29-NOV-2001.
 XX
 PF 18-MAY-2001; 2001MO-US16450.
 XX
 PR 19-MAY-2000; 2000US-205515P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Birse CE, Rosen CA;
 XX
 DR WPI, 2002-122018/16.
 DR N-PSDB; ABL89762.
 XX
 PT Novel 1405 isolated polypeptides, useful for diagnosis, treatment and
 PT prevention of neural, immune system, muscular, reproductive, and
 PT gastrointestinal, pulmonary, cardiovascular, renal and proliferative
 PT disorders -
 XX
 PS Claim 11; SEQ ID NO 1729; 2081pp + Sequence Listing; English.
 XX
 CC The invention relates to novel genes (ABL89449-ABL90853) and proteins
 CC (ABB89040-ABB90444) useful for preventing, treating or ameliorating
 CC medical conditions e.g. by protein or gene therapy. The genes are
 CC isolated from a range of human tissues disclosed in the specification.
 CC The nucleic acids, proteins, antibodies and (ant)agonists are useful
 CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
 CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
 CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
 CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
 CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
 CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
 CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;
 CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
 CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
 CC and parasitic infections.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 190 AA;

Percent Similarity: 100.00%
 Best Local Similarity: 100.00%
 Query Match: 49.10%
 DB: 23
 Gaps: 0
 US-09-896-522-3 (1-834) x ABB89353 (1-190)
 QY 100 AGCGGGAAGTCGACCGTGTGTGAGATCATGAGATTGCTGGACAGAAAGAGTGGAA 159
 |||||
 Db 34 SerGLySerThrValCysGLySILeMeGLInleuGLyGLInengIuValGLu 53
 |||||
 QY 160 CAGCGGCAAGAGAGGTGTCACTCTGAGCCAGAGACAGGTTCTACAGGTCTGACGGCA 219
 |||||
 Db 54 GlnArgGLnArgGLyValIleuSerGlnAspArgPheTyrValIleuThrAla 73
 |||||
 QY 220 GAGCAGAAAGCCAGGCTTTGAAAGACGTACATTTTGACCATTCAGATGCTTTGAT 279
 |||||
 Db 74 GLuGLInLysAlaLysAlaLeuLysGLInTyrAsnPhaAspHisProAspAlaPheAsp 93
 |||||
 QY 280 AATGATTGATGACAGCACTCTGAAGAACATCTGTGAGGCAAAAGGTTGAGTGGCG 339
 |||||
 Db 94 AsnAspLeuMetHisArgThrLeuLysAsnIleValGLuGLyThrValGLuValPro 113
 |||||
 QY 340 ACCTATGATTTTGTGACACACTCAAGTTACAGACAGACAGCGTGTCTACCTGCGGAC 399
 |||||
 Db 114 ThrTyrAspPheValThrHisSerArgLeuProGLuThrValValTyrProAlaAsp 133
 |||||
 QY 400 GTGTGTTCTGTGTGAGGCACTTTGTGTCTACAGCCAGAGATCCGGGACATGTTCCAC 459
 |||||
 Db 134 ValValLeuPheGLuGLuIleLeuValPheTyrSerGLInLylLeArgAspMetPheHis 153
 |||||
 QY 460 CAGCGGCTCTGTGTGACACAGCACTCCGACGTCCAGCTGTCTGAGGA 507
 |||||
 Db 154 LeuArgLeuPheValAspThrAspSerAspValArgLeuSerArgArg 169
 |||||
 RESULT 8
 ABB89328
 ID ABB89328 standard; Protein; 120 AA.
 XX
 AC ABB89328;
 XX
 DT 21-OCT-2002 (first entry)
 XX
 DE Human MDDT protein Incyte ID No: LI:235557.12.orf2:2001JAN12.
 XX
 XX Human; molecule for disease detection and treatment; MDDT; cancer;
 KM cell proliferative disorder; arteriosclerosis; cirrhosis; asthma;
 KM autoimmune disorder; inflammatory disorder; Crohn's disease;
 KM multiple sclerosis; cytoskeletal; antiarteriosclerotic; antiinflammatory;
 KM hepatotropic; immunosuppressive; antiasthmatic.
 XX
 OS Homo sapiens.
 XX
 PN WO200255738-A2.
 XX
 PD 18-UTL-2002.
 XX
 PF 09-JAN-2002; 2002MO-US01008.
 XX
 PR 12-JAN-2001; 2001US-261622P.
 PR 16-JAN-2001; 2001US-261655P.
 PR 17-JAN-2001; 2001US-262208P.
 PR 17-JAN-2001; 2001US-262209P.
 PR 17-JAN-2001; 2001US-262326P.
 PR 19-JAN-2001; 2001US-263063P.
 PR 19-JAN-2001; 2001US-263065P.
 PR 19-JAN-2001; 2001US-263329P.
 XX
 PA (INCY-) INCYTE GENOMICS INC.
 XX
 PI Panzer SR, Lincoln SE, Altus CM, Dufour GE, Hillman UT, Jones AL;
 PI Dam TC, Liu TF, Harris B, Flores V, Dafio A, Marwaha R, Chen AJ;
 PI Chang SC, Gerstin EH, Peralta CH, David MH, Lewis SA;

XX WPI: 2002-590679/63.
DR N-PSDB: ABS51801.
XX
PT New disease detection and treatment molecule (MDT) polynucleotides and
PT polypeptides, useful in diagnosing, studying, preventing or treating
PT diseases associated with MDT expression, e.g. autoimmune or
PT inflammatory disorders
XX
PS Claim 27, Page 120; 129pp; English.
XX
CC The present invention relates to the isolation of novel human
CC molecules for disease detection and treatment (MDT), and the
CC polynucleotide sequences (mdt) encoding them. The MDT polypeptides
CC may be used to screen for molecules that bind to, or are bound by the
CC encoded polypeptides, and to develop a transcript image of a tissue or
CC cell type. Probes comprising at least 20 nucleotides of the mdt
CC polynucleotide may be used to assess the toxicity of a test compound.
CC The MDT polypeptides and mdt polynucleotides are useful in the
CC diagnosis, study, prevention and treatment of diseases associated with
CC the expression of molecules for disease detection and treatment. Such
CC disorders include cell proliferative disorders (e.g. arteriosclerosis,
CC cirrhosis, or cancers), and autoimmune/inflammatory disorders
CC (e.g. asthma, Crohn's disease, or multiple sclerosis). The mdt
CC polynucleotides may also be used as molecule markers, in microarrays,
CC and in somatic or germline gene therapy. ABG70306-ABG70341 represent
CC the MDT proteins of the invention.
XX

Sequence 120 AA;

Alignment Scores:
Pred. No.: 1,59e-104 Length: 120
Score: 112.00 Matches: 112
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 40.43% Indels: 0
DB: 23 Gaps: 0

US-09-896-522-3 (1-834) x ABG70328 (1-120)

QY 268 GATGCTTGTATGATTTGATGACAGACCTGTGAAGAACATCTGAGGGCAAAACG 327
DB 5 AspIaIahepaaspaasplemechisaArgThreulysanilleValIGluGlyLysThr 24
QY 328 GTGAGGAGTGGCGACCTATGATTTGTGACACACTCAAGTTTACAGAGACCAAGCGTGC 387
DB 25 ValGluValProthrTyraAspPheValThrHisSerArgLeuProGluThrValVal 44
QY 388 TACCCTGCGAGCGTGTCTGTTGAGGGCATCTTGTGTTCTACAGCCAGAGATCCGG 447
DB 45 TyrProAlaAspValValIleuPheGluGlyIleuValPheTyrSerGlnGluIleArg 64
QY 448 GACATGTTCCACCGCGCTCTTGTGAGACCGACCTCCGACCGCTGTCTCGAAGA 507
DB 65 AspMetPheHisIleuArgLeuPheValAspThrAspSerAspValArgLeuSerArgArg 84
QY 508 GTTCTCCGGAGCGTCCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 567
DB 85 ValIleuArgAspValAlaArgArgGlyArgAspLeuGluIleuIleuThrGlnTyrThrThr 104
QY 568 TTGCTGAAGCCGCGCTTCTGAGAGAGTTCGCTCGCG 603
DB 105 PheValIysProAlaPheGluGluPheCysIleuPro 116

RESULT 9

AAOI4413
ID AAOI4413 standard; Protein; 277 AA.
XX
AC AAOI4413;
XX
DT 02-MAY-2002 (first entry)
XX
DE Protein relating to a human uridine kinase (UDK) of the invention.

XX Human; uridine kinase; diagnostic assay; mutation detection; UDK;
KM probe; chromosome localisation study; tissue expression; gene therapy;
KM antibody; vaccine; human ovarian cancer; immunological disorder;
KM human colon carcinoma; immunogen.
XX
OS Unidentified.
XX
PN WO200172963-A2.
XX
PD 04-OCT-2001.
XX
PF 27-MAR-2001; 2001WO-US09663.
XX
PR 27-MAR-2000; 2000US-0536647.
XX
PA (SMIK) SMITHKLINE BEECHAM CORP.
XX
PI Ho YS, Johnson RK;
XX
DR WPI: 2001-626259/72.
XX
PT Novel human uridine kinase polypeptides useful for treating cancers,
PT and to identify agonists and antagonists of the polypeptide useful for
PT treating conditions associated with uridine kinase imbalance
XX
PS Disclosure; Page 23; 31pp; English.

CC The invention relates to newly identified human uridine kinase (UDK)
CC polypeptides and polynucleotides and methods for producing such
CC polypeptides by recombinant techniques. Also disclosed in the invention
CC are methods for utilising uridine kinase polypeptides and polynucleotides
CC in diagnostic assays. The polynucleotides and polypeptides of the
CC invention may be used as diagnostic reagents by detecting mutations in an
CC associated gene. An array of oligonucleotide probes comprising the
CC uridine kinase polynucleotide sequence or fragments thereof can be
CC constructed to conduct efficient screening of genetic mutations, for
CC example. Detection of abnormally decreased or increased levels of
CC polypeptide or mRNA expression may also be used for diagnosing or
CC determining susceptibility of a subject to a disease of the invention.
CC The polynucleotide sequences of the invention can be used for chromosome
CC localisation studies and tissue expression studies. The polypeptides of
CC the invention or fragments thereof may be used as immunogens to produce
CC antibodies. These antibodies may be employed to isolate or identify
CC clones expressing the polypeptide. The polypeptides and polynucleotides
CC of the invention can be used as a vaccine or in gene therapy to treat
CC diseases such as human ovarian cancer, human colon carcinomas, and
CC immunological disorders. This sequence represents the protein relating to
CC a human uridine kinase (UDK) of the invention.
CC NOTE: The present sequence is stated as being the same as that shown as
CC SEQ ID NO: 2 in the sequence listing of the specification. However, the
CC sequences differ.

Sequence 277 AA;

Alignment Scores:
Pred. No.: 1,42e-36 Length: 277
Score: 45.00 Matches: 45
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 16.25% Indels: 0
DB: 22 Gaps: 0

US-09-896-522-3 (1-834) x AAOI4413 (1-277)

QY 298 ACTCTGAAGAACATCTGTGAGGGGCAAAACGTTGAGAGTCCGACCTATGATTTGTGACA 357
DB 100 ThrIleuYAsnIleValIGluGlyLysThrValIGluValProThrTyraAspPheValThr 119
QY 358 CACTCAAGTTTACAGAGACCAAGCGTGTCTACCTCGGAGAGTGGTCTGTTGAGGCG 417
DB 120 HisSerArgLeuProGluThrThrValValTyrProAlaAspValValIleuPheGluGly 139

Oy	418	ATCTTGTTGTTCTAC	4332
Db	140	IleuValPheTyr	144
RESULT 10			
ID	AA038694		
ID	AA038694	standard; Protein; 261 AA.	
XX	AA038694;		
XX			
DT	22-OCT-2001	(first entry)	
XX			
DE	Human polypeptide SEQ ID NO 1839.		
XX			
KW	Human; neurotropic; immunosuppressant; cytostatic; gene therapy; cancer;		
KW	peripheral nervous system; neuropathy; central nervous system; CNS;		
KW	Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;		
KW	amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;		
KW	chemokine; thrombolytic; drug screening; arthritis; inflammation;		
KW	leukaemia.		
XX			
OS	Homo sapiens.		
XX			
PN	WO200153312-A1.		
XX			
PD	26-JUL-2001.		
XX			
PF	26-DEC-2000; 2000MO-US34263.		
XX			
PR	21-JAN-2000; 2000US-0488725.		
PR	25-APR-2000; 2000US-0552317.		
PR	09-JUL-2000; 2000US-0598042.		
PR	19-JUL-2000; 2000US-0620312.		
PR	03-AUG-2000; 2000US-0653450.		
PR	14-SEP-2000; 2000US-0662191.		
PR	19-OCT-2000; 2000US-0693036.		
PR	29-NOV-2000; 2000US-0727344.		
XX			
PA	(HYSE-) HYSEQ INC.		
PI	Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;		
PI	Mang J, Wang Z, Wehman T, Xu C, Xue AJ, Yang Y, Zhang J;		
PI	Zhao QA, Zhou P, Goodrich R, Drmanac RT;		
XX			
DR	WPI: 2001-442253/47.		
DR	N-PSDB; AAI57850.		
XX			
PT	Novel nucleic acids and polypeptides, useful for treating disorders		
PT	such as central nervous system injuries -		
XX			
PS	Example 3; SEQ ID NO 1839; 10078pp; English.		
XX			
CC	The invention relates to human nucleic acids (AA157798-AA161369) and		
CC	the encoded polypeptides (AA038642-AA042213) with neurotropic,		
CC	immunosuppressant and cytostatic activity. The polynucleotides are useful		
CC	in gene therapy. A composition containing a polypeptide or polynucleotide		
CC	of the invention may be used to treat diseases of the peripheral nervous		
CC	system, such as peripheral nervous injuries, peripheral neuropathy and		
CC	localised neuropathies and central nervous system diseases, such as		
CC	Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic		
CC	lateral sclerosis, and Shy-Drager Syndrome. Other uses include the		
CC	utilisation of the activities such as: Immune system suppression,		
CC	Activin/inhibin activity, chemotactic/chemokine activity, haemostatic		
CC	and thrombolytic activity, cancer diagnosis and therapy, drug screening,		
CC	assays for receptor activity, arthritis and inflammation, leukaemias and		
CC	C.N.S disorders.		
CC	Note: The sequence data for this patent did not form part of the printed		
CC	specification.		
XX			
XX	Sequence 261 AA;		
XX			

Alignment Scores: 2.84e-16

Length: 261

Prod. No.:

Score:	25.00	Matches:	25
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatch:	0
Query Match:	9.03%	Indels:	0
DB:	22	Gaps:	0
US-09-896-522-3 (1-834) x AAM38694 (1-261)			
QY	565	ACCTTCGTGAAGCCGCTTCGAGGAGTTCTGCCTGCCGACAAAGAGTATGCCGATGTG	624
Db	187	ThrpheValIysProAlaPheGluIuIbheCysIeuProThrIyIeSyrAlaSepVal	206
QY	625	ATCATCCGACGAGCA	639
Db	207	IleIleProArgGly	211
RESULT 11			
AAB73494			
ID	AAB73494	standard; Protein; 261 AA.	
XX	AAB73494;		
DT	31-JUL-2001	(first entry)	
XX	Human transferrase HTFS-1, SEQ ID NO:1.		
XX	Human transferrase; HTFS; agonist; antagonist; cellular signalling;		
KW	proliferation; cell proliferative disorder; immune disorder;		
KW	atherosclerosis; hepatitis; psoriasis; cancer; tumour;		
KW	inflammation; AIDS; Addison's disease; allergy; asthma; anaemia;		
KW	cirrhosis; Crohn's disease; atopic dermatitis; diabetes mellitus;		
KW	multiple sclerosis; rheumatoid arthritis; pancreatitis;		
KW	systemic lupus erythematosus; thrombocytopenia; ulcerative colitis;		
KW	haemodialysis; extracorporeal circulation; trauma; transgenic animal;		
XX	gene therapy; drug screening.		
XX			
OS	Homo sapiens.		
XX			
PN	WO20032888-A2.		
XX			
PD	10-MAY-2001.		
XX			
XX	02-NOV-2000; 2000WO-US30485.		
PF			
XX			
PR	04-NOV-1999; 99US-0163595.		
XX			
PA	(INCYTE) INCYTE GENOMICS INC.		
XX			
PI	Tang YT, Yue H, Hillman JL, Lal P, Bandman O, Patterson C;		
PI	Shih IL, Azimzal Y, Lu DM, Baughn MR;		
XX			
DR	MP1: 2001-328796/34.		
XX	N-P8DB; AAM23801.		
PT	Human transferrase polypeptides and polynucleotides useful for		
PT	diagnosis, prevention and treatment of cell proliferative and immune		
PT	system disorders and for identifying agonists and antagonists -		
XX			
XX			
P8	Claim 1; Page 103-104; 157pp; English.		
XX			
CC	Sequences AAB73494-AAB73535 represent novel human transferrase proteins		
CC	HTFS-1 to HTFS-42, and sequences AAM23801-AAM23842 represent cDNAs		
CC	encoding them. The proteins play important roles in the regulation of		
CC	cellular signalling and proliferation. The HTFS proteins are useful for		
CC	screening compounds for their effectiveness as agonists or antagonists of		
CC	transferrase activity, or for compounds that specifically bind to an HTFS		
CC	protein or which modulates the activity of an HTFS protein.		
CC	Pharmaceutical compositions comprising an HTFS protein, HTFS		
CC	agonist or antagonist, or genetic construct encoding an HTFS		
CC	protein are useful for treating a disease or condition associated		
CC	with decreased or increased expression of functional HTFS. Disorders		
CC	which may be treated using such compositions include cell proliferative		
CC	disorders and immune disorders. For example, diseases which may be		

CC treated include atherosclerosis, hepatitis, psoriasis, cancers (including
 CC breast, bladder, bone marrow, brain and uterus cancer), inflammation,
 CC AIDS, Addison's disease, allergies, asthma, anemia, cirrhosis, Crohn's
 CC disease, atopic dermatitis, diabetes mellitus, multiple sclerosis,
 CC rheumatoid arthritis, pancreatitis, systemic lupus erythematosus,
 CC thrombocytopenia, and ulcerative colitis. They may also be used to treat
 CC complications of cancer, haemodialysis, extracorporeal circulation,
 CC trauma and haemopoietic cancer, including lymphoma, leukaemia and
 CC myeloma. Polynucleotides encoding HTPS proteins are useful for creating
 CC transgenic animals to model human diseases, for diagnostic purposes and
 CC to generate hybridisation probes useful in mapping the naturally
 CC occurring genomic sequences. HTPS, and its catalytic or immunogenic
 CC fragments are useful for screening libraries of compounds in a variety of
 CC drug screening techniques. Antibodies which specifically bind HTPS may be
 CC used for the diagnosis of disorders associated with the expression of
 CC HTPS, or in assays to monitor patients being treated with HTPS or
 CC agonists, antagonists or inhibitors of HTPS. The present sequence
 CC represents an HTPS protein of the invention.

XX Sequence 261 AA;

Alignment Scores:

Pred. No.:	2,846-16	Length:	261
Score:	25.00	Matches:	25
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	9.03%	Indels:	0
DB:	22	Gaps:	0

US-09-896-522-3 (1-834) x AAB73494 (1-261)

QY 565 ACCTTGGAAGCCGCTTCGAGAGTTCTGCTGCCGACAAAGATATGCCATGTG 624
 |||
 Db 187 ThrPheVallylsProAlaPheGluGluPheCysLeuProThrlySTyAlaAspVal 206

QY 625 ATCATCCACGAGGA 639
 |||
 Db 207 IleIleProArgly 211

RESULT 12

AAB56582
 ID AAB56582 standard; Protein: 337 AA.

XX AAB56582;

DT 13-MAR-2001 (first entry)

DE Human prostate cancer antigen protein sequence SEQ ID NO:1160.

XX Human; prostate cancer; prostate cancer antigen; detection; diagnosis;
 KW neuroprotective; cytoskeletal; cardiovascular; immunomodulatory; muscular;
 KW vulnery; gastrointestinal; nephrotoxic; antiinfective; gynaecological;
 KW antibacterial; gene therapy; neural; immune; reproductive; renal;
 KW gastrointestinal; pulmonary; cardiovascular; proliferative disorder;
 KW wound; infectious disease.

XX Homo sapiens.

PN WO200055174-A1.

PD 21-SEP-2000.

PF 08-MAR-2000; 2000WO-US05988.

PR 12-MAR-1999; 99US-0124270.

PA (HUMA-) HUMAN GENOME SCI INC.

PA (ROSE/) ROSEN C A.

PI Rosen CA, Ruben SM;

XX WPI, 2000-587513/55.
 DR N-PSDB; AAF15785.

XX Prostate cancer associated gene sequences, referred to as prostate
 PT cancer antigens, useful for treatment, prevention, and diagnosis of
 PT disorders such as prostate cancer -

XX Claim 11; Page 1566-1567; 2338pp; English.

XX AAF1566 to AAF16505 encode the human prostate cancer associated
 CC proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.
 CC The prostate cancer antigens can have neuroprotective, cytoskeletal,
 CC cardiovascular, immunomodulatory, muscular, vulnery, gastrointestinal,
 CC nephrotoxic, antiinfective, gynaecological and antibacterial activities,
 CC and can be used in gene therapy. The prostate cancer antigen
 CC polynucleotides may be used for detection of prostate cancer, chromosome
 CC identification, as chromosome markers, and for numerous other diagnostic
 CC or research purposes. The prostate cancer antigens may be used to treat
 CC disorders such as neural, immune, muscular, reproductive,
 CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
 CC disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to
 CC AAB57303 represent sequences used in the exemplification of the present
 CC invention.

XX Sequence 337 AA;

Pred. No.:	2,766-16	Length:	337
Score:	25.00	Matches:	25
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	9.03%	Indels:	0
DB:	21	Gaps:	0

US-09-896-522-3 (1-834) x AAB56582 (1-337)

QY 565 ACCTTGGAAGCCGCTTCGAGAGTTCTGCTGCCGACAAAGATATGCCATGTG 624
 |||
 Db 263 ThrPheVallylsProAlaPheGluGluPheCysLeuProThrlySTyAlaAspVal 282

QY 625 ATCATCCACGAGGA 639
 |||
 Db 283 IleIleProArgly 287

RESULT 13

ABP41393
 ID ABP41393 standard; Protein: 337 AA.

XX ABP41393;

DT 22-AUG-2002 (first entry)

DE Human ovarian antigen HOELP29, SEQ ID NO:2525.

XX Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;
 KW ovarian cancer; breast cancer; tumour; reproductive system disorder;
 KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
 KW PCOS; ovarian cyst; dysmenorrhoea; endocrine disorder; infection;
 KW inflammatory condition; immune disorder; blood disorder;
 KW cardiovascular disorder; respiratory disorder; neurological disorder;
 KW gastrointestinal disorder; urinary system disorder; drug screening;
 KW gene therapy; chromosome mapping; forensic analysis;
 KW antibody preparation; cytoskeletal; immunomodulatory; neuroprotective;
 KW antiinflammatory; gynaecological; reproductive; chromosome 1p32.

XX Homo sapiens.

PN WO200200677-A1.

PD 03-JAN-2002.

PF 07-JUN-2001; 2001WO-US18569.

PR 07-JUN-2000; 2000US-209467P.


```

XX 26-MAR-2002 (first entry)
DT
XX Drosophila melanogaster polypeptide SEQ ID NO 13713.
DE
XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX Drosophila melanogaster.
OS
XX WO200171042-A2.
FN
XX 27-SEP-2001.
PD
XX 23-MAR-2001; 2001WO-US09231.
PF
XX 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
XX (PEKE ) PE CORP NY.
PA
XX Venter JC, Adams M, Li PWD, Myers EW;
PI
XX WPI; 2001-656860/75.
DR N-PSDB; ABL06410.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
XX Disclosure; SEQ ID NO 13713; 21pp + Sequence Listing; English.
PS
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
CC sequences (AB101840-AB16175) and the encoded proteins
CC (AB157737-AB172072).
CC
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX
SQ Sequence 260 AA;

```

Alignment Scores:

Pred. No.:	4e-06	Length:	260
Score:	15.00	Matches:	15
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	5.42%	Indels:	0
DB:	22	Gaps:	0

US-09-896-522-3 (1-834) x ABB62307 (1-260)

```

OY 388 TACCGGCGGAGCGTCTGTTGAGGGGACCTTGNGTCTAC 432
    |||||||
Db 135 TyrProAlaAspValValLeuPheGluCylIleLeuValPheIyr 149

```

Search completed: November 25, 2003, 07:53:19
 Job time : 34.5012 secs

Alignment Scores:

Pred. No.: 2,58e-80 Length: 201
 Score: 1020.50 Matches: 200
 Percent Similarity: 85.47% Conservative: 0
 Best Local Similarity: 85.47% Mismatches: 1
 Query Match: 67.40% Indels: 33
 DB: 4 Gaps: 1

US-09-896-522-3 (1-834) x Q96BU0 (1-201)

QY 1 ATGGCTTCGGCGGAGGCGGAAGACTGCGAGACCCCGCGCGGACCGCTCCGAC 60
 Db 1 MetlaSerlaaglygluAspCysgluSerProAlaProgluAlaAspArgProHis 20
 QY 61 CAGCGGCCCTCTGATAGGGGTGAGCGCGGCACTGCGACGGGAAAGTCAGCTGTGT 120
 Db 21 GlnArgProPheLeuIleGlyValSerGlyGlyThrAlaSerGlyLysSerThrValCys 40
 QY 121 GAGAAAGTCAATGAGTGTCTGGAGCAGAAAGAGGTGGAACGCGGACGGGAGGTGTC 180
 Db 41 GlnuylleuMetGluLeuGluGlnAsnGluValGluGlnArgGlnArgLysValVal 60
 QY 181 ATCTGAGCCAGGACAGGTTCTCAAGGTCCTGACGCGACAGCAAGCCAGCCCTTG 240
 Db 61 IleuSerGlnAspArgPheTyrLysValLeuThrAlaGluGlnLysAlaLysAlaLeu 80
 QY 241 AAAAGACATGCAATTTTGCACATCCAGATCCCTTGTATATGATTTGATGACAGGACT 300
 Db 81 LysGlyGlnIleuPheAspHisProAspAlaPheAspAspLeuMetHisArgThr 100
 QY 301 CTGAAGAACATCGTGAAGGCGGCAAAACGGTGAAGGCGGACCTATGATTTTGTGACAC 360
 Db 101 LeuLysAsnIleValIleGluGlyLysThrValGluValProThrTyrAspPheValThrHis 120
 QY 361 TCAAGTTACGAGACGACCGGCGTGTCTACCTCGCGACGCGTGTCTGTTGAGGCGATC 420
 Db 121 SerArgLeuProGluThrThrValValTyrProAlaAspValValLeuPheGluGlyLe 140
 QY 421 TTGGTGTCTACAGCCAGGAGATCCGGGACATGTTCCACCTGCGCCTCTTGTGAGACAC 480
 Db 141 LeuValPheTyrSerGlnGluIleArgAspMetPheHisLeuArgLeuPheValAspThr 160
 QY 481 GACTCCGACGTCAGGCGTGTCTGAAGAGTTCTCCGGACGCGCCGACGAGGAGGACCTG 540
 Db 161 AspSerAspValArgLeuSerArg--Arg----- 169
 QY 541 GAGCAGATTCTGACGACGACGACCACTTCGTGAAGCGCGCTTCGAGAGTTCTGCTG 600
 Db 169 ----- 169
 QY 601 CCGGCAAGAAGTATGCCGATGTGATCATCCGACGAGAGTGGACATATGTGTCATC 660
 Db 170 --AspLysGluValCysArgCysAspHisProThrArgSerGlyGlnTyrCysHisG 189
 QY 661 AACCTGATCGTGACGACATCCAGACATTCCTGATGAG 698
 Db 189 InProAspArgAlaAlaHisProGlyHisSerGluTyr 201

RESULT 2
 Q8C476 PRELIMINARY; PRT; 261 AA.
 AC Q8C476;
 DT 01-MAR-2003 (TReMBLrel. 23, Created)
 DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
 DE Uridine-cytidine kinase 2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J;

RX MEDLINE=22354683; PubMed=12466851;
 RA The FANTOM Consortium,
 RA the RIKEN Genome Exploration Research Group Phase I & II Team,
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs."
 RL Nature 420:563-573 (2002).
 DR EMBL; AK082837; BAC38646.1;
 SQ SEQUENCE 261 AA; 29438 MW; 780AA3C298AA0153 CRC64;

Alignment Scores:

Pred. No.: 1.41e-72 Length: 261
 Score: 932.00 Matches: 187
 Percent Similarity: 79.04% Conservative: 28
 Best Local Similarity: 68.75% Mismatches: 41
 Query Match: 61.56% Indels: 16
 DB: 11 Gaps: 4

US-09-896-522-3 (1-834) x Q8C476 (1-261)

QY 22 GACTGGAGAGAGCCCGCGCGGAGCGGACCGTCCGAC---CAGCGCCCTTCTGATA 78
 Db 4 AspSerGluGlnThrLeuGlnAsnHisGlnGlnProAsnGlyGlyGluProPheLeuIle 23
 QY 79 GGGGTGAGCGCGGACCTGCGACGCGGAGTGCACCGTGTGAGAAAGATCATGAGATTG 138
 Db 24 GlyValSerGlyGlyThrAlaSerGlyLysSerSerValCysAlaLysIleValGlnLeu 43
 QY 139 CTGGACAGAAAGAGGTGAAACGCGGACGCGGAGGTGTCATCTGACCGACAGACAG 198
 Db 44 LeuGlyGlnAsnGluValAspTyrHisGlnLysGlnValIleLeuSerGlnAspSer 63
 QY 199 TTCTACAGGTCCTGACGCGGAGCGGAAAGCGGACCGTGAAGACATGATTTT 258
 Db 64 PheTyrArgValLeuThrSerGluGlnLysAlaLysAlaLeuLysGlyLysPheAspLeu 83
 QY 259 GACCATCCAGATGCTTTGATATGATTTGATGACAGACTCTGAAGACATCTGAGAG 318
 Db 84 AspHisProAspAlaPheAspAsnGluLeuIlePheLysThrLeuLysGluIleThrGlu 103
 QY 319 GGCAGAAACGGTGAAGGTCCGACCTGATGATTTTGTGACACATCAAGATTACAGACAC 378
 Db 104 GlyLysThrValGlnIleProValTyrAspPheValSerHisSerArgLysGluGlnThr 123
 QY 379 ACGTGTGTACCTCGGACGCGTGTCTGTTGAGGAGATCTGTGTTCTTCAAGCCAG 438
 Db 124 ValThrIleTyrProAlaAspValValLeuPheGluGlyLeuAlaPheTyrSerGln 143
 QY 439 GAGATCCGGACATGTTCCACCTGCGCCTCTTGTGAGACGCACTCCGACGTCAAGCTG 498
 Db 144 GluValArgAspLeuPheGlnMetLysLeuPheValAspThrAspAlaAspThrArgLeu 163
 QY 499 TCTGAAAGATTCTCCGGAAGTG---CGCCGAGGAGGAGCCTGAGACAGATTCTGACG 555
 Db 164 SerArgArgValLeuArgAspLysSerGluArgGlyArgAspLeuGlnIleLeuSer 183
 QY 556 CAGTACACACACTTGTGAAGCGGCTTCGAGGATTCGTGCTGCGCAAGAAGATAT 615
 Db 184 GlnTyrIleThrPheValLysProAlaPheGluGluPheCysLeuProThrLysLysTyr 203
 QY 616 GCCGATGTGATCATCCACGAGAGTGAACAATATGTTGCCATCACTGATCGTGAC 675
 Db 204 AlaAspValIleIleProArgGlyAlaAspAsnLeuValAlaIleAsnLeuIleValGln 223
 QY 676 CACATCCAGACATTTGTAATGTGATCATCTGCAAAATGGACCGGAGAGGTCCAAATGG 735
 Db 224 HisIleGlnAspIleLeuAsnGly-----Gly 232
 QY 736 CGAGGTCAAGGAGACCTTTTCTGACGAGGAGCACACCTGGAGATGCTGACTGTGGC 795
 Db 233 LeuSerLysArgGlnThr-----AsnGlyTyrPheAsnGlyTyrThrProSerArg 249
 QY 796 AACGCTCACTTTGAGTTCAGACAGACAGCCAC 831

Db 250 LysarglnalaserGluSerSerSexArgProHis 261

RESULT 3

Q8MRJ1 PRELIMINARY; PRT; 260 AA.

AC Q8MRJ1; 260 AA.

DT 01-OCT-2002 (TREMBlrel. 22, Created)

DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)

DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)

DE L013909P.

GN C6364.

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI_TaxID=7227;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Berkley;

RA Stapleton M., Brokstein P., Hong L., Aghayani A., Carlson J.,

RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,

RA George R., Gonzalez W., Gharin H., Krommiller B., Li P., Liao G.,

RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,

RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,

RA Ceiniker S.;

RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL: AY119583; AAM50237.1; -

DR FlyBase; FBgn0039179; CG6364.

DR InterPro; IPR006083; PCK_URK.

DR InterPro; IPR007664; Uridine_kin.

DR Pfam; PF00485; PRK; 1.

DR PRINTS; PR00988; URIDINKINASE.

DR SEQUENCE 260 AA; 29223 MW; 03CAA6DCA04A1655 CRC64;

Alignment Scores:

Pred. No.: 3.03e-55 Length: 260

Score: 733.00 Matches: 142

Percent Similarity: 82.46% Conservative: 32

Best Local Similarity: 67.30% Mismatches: 35

Query Match: 48.41% Indels: 2

DB: Gaps: 2

US-09-896-522-3 (1-834) x Q8MRJ1 (1-260)

QY 67 CCCTCTCTGTAAGGGGAGGAGCGGCGGCACTGCCAGCGGAAGTGCCTGTGTGAGAG 126

Db 27 PpPheLeuIleGlyValAlaGlyGlyThrAlaSerGlyYsserThrValCysLysLys 46

QY 127 ATCATGAGTGTGCGGAG 186

Db 47 IleMeGluGlnLeuGlyGlnAlaGlnMeCaspHisThrGlnMglnValIleValSerIle 66

QY 187 AGCCAGAGAGAGGTTCTTCAAGAGTCTTACGCGAGAGAGAGAGAGAGAGAGAGAGAG 246

Db 67 SerGlnAspSerPheTyrArgGlnLeuThrProAlaGlnLysAlaLysGlnLysGly 86

QY 247 CAGTACATTTTGAACCATCCAGATGCTTGTATATATATTGATGCACAGAGCTTGAAG 306

Db 87 LeuPheAsnIlePheAspHisProAspAlaPheAsnIleLeuLeuMetTyrSerThrLeuGln 106

QY 307 AACATCGTGGAGGCAAAACGTTGAGAGTCCGACCATGATTTTGGACACACTCA--- 363

Db 107 AsnIleLeuLysGlyHisLysValGlnIleProSerTyrAspTyrArgThrAsnSerLeu 126

QY 364 AGGTTACAG 423

Db 127 AspPheGluAsnValLeuValIleTyrProAlaAspValValLeuPheGluGlyIleLeu 146

QY 424 GTGTTCACAG 483

Db 147 ValPheTyrPheProLysIleArgGlnLeuPheIleMetLysLeuPheValAspThrAsp 166

QY 484 TCCAGAGTCAAGGCTGTCTCGAAGAGTTCCTCCGGAGAGCTG---CGCCAGAGAGAGAG 540

Db 167 ProAspThrArgLeuAlaArgValProArgAspIleAsnIleLysGlnArgAlaGlyLeu 186

QY 541 GAGCAGATTCTGACGAGATACACACCTTGTGTAAGCGGCGCTTGCAGAGAGTTCGCTG 600

Db 187 AspAlaValLeuThrGlnTyrMetThrPheValLysProAlaPheGluGlnPheCysSer 206

QY 601 CCGACAAAGAGAGATGCCGATGATCATCCGACGAGAGAGAGAGAGAGAGAGAGAGAG 660

Db 207 ProThrLysPheAsnValIleIleProArgGlyAlaAspAsnThrValAlaIle 226

QY 661 AACCTGATCGTGCAGACATCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 693

Db 227 AspLeuIleValIleHisIleGlyGlnIleLeu 237

RESULT 4

Q9FKS0 PRELIMINARY; PRT; 486 AA.

AC Q9FKS0;

DT 01-MAR-2001 (TREMBlrel. 16, Created)

DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)

DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)

DE Uridine kinase-like protein.

GN AT5G40870.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC Eurosid II; Brassicales; Brassicaceae; Arabidopsis.

OX NCBI_TaxID=3702;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Columbia;

RA MEDLINE=98344145; PubMed=9679202;

RA Kaneo T., Kotani H., Nakamura Y., Sato S., Asamizu E., Miyajima N.,

RA Tabata S.;

RT "Structural analysis of Arabidopsis thaliana chromosome 5. V. Sequence

RT features of the regions of 1,381,565 bp covered by twenty one

RL physically assigned p1 and TAC clones.";

RL DNA Res. 5:131-145(1998).

RN [2]

RP SEQUENCE FROM N.A.

RA Yamada K., Chan M.M., Chang C.H., Dale J.M., Hsuan Y.W., Lee J.M.,

RA Onodera C.S., Quach H.L., Tang C.C., Toriumi M., Wong C., Wu H.C.,

RA Yu G., Yuan S., Chen H., Cheuk R., Jones T., Kim C.J., Nguyen M.,

RA Palm C.U., Shim P., Southwick A., Tripp M.G., Wu T., Davis R.W.,

RA Ecker J.R., Theologis A.;

RT "Arabidopsis Open Reading Frame (ORF) Clones.";

RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL: AB011477; BAB11349.1; -

DR EMBL: BT002336; AAN86169.1; -

DR HSSP; Q26996; 1BD3.

DR InterPro; IPR006083; PRK_URK.

DR InterPro; IPR000836; PRTtransferase.

DR Pfam; PF00156; Priboyltran_1.

DR Pfam; PF00485; PRK; 1

DR PRINTS; PR00988; URIDINKINASE.

DR TIGRPFAM; TIGR00235; udk; 1.

KW Kinase; Transferase.

SO SEQUENCE 486 AA; 54430 MW; 50DA1CE89346FB54 CRC64;

Alignment Scores:

Pred. No.: 5.09e-41 Length: 486

Score: 570.50 Matches: 119

Percent Similarity: 66.40% Conservative: 49

Best Local Similarity: 47.04% Mismatches: 68

Query Match: 37.68% Indels: 17

DB: Gaps: 7

US-09-896-522-3 (1-834) x Q9FKS0 (1-486)

QY 31 AGCCCGCGCCGAG 84

```

Db      48 SerSerSerProSerSerSerSerProGluAlaProLysGlnProPheIleIleGlyVal 67
Qy      85 AGCGCGGCGACTGCCAGCGGGAGATCGACCGTGTGTGAAGAAATCATGAGACTTCTGGGA 144
Db      68 SerGlyGlyThrAlaSerGlyLysThrValCysAspMetIleIleGlnGlnLeu--- 86
Qy      145 CAGAACGAGGTGGAACAGCGCGAGCGGAGGTGGTCACTCCGAGACCGAGACAGGTTCTAC 204
Db      87 -----HisAspHisArgValAlaLeuValAlaGlnAspSerPheTyr 100
Qy      205 AAGGCTCTGAGCGAGACAGAGAGCGGACCGCTTGAAGAGCACTCAATTTTGACCAT 264
Db      101 ArgGlyLeuThrSerGlnGlnLeuGlnArgVal-----GlnGlyTyrAsnPheAspHis 118
Qy      265 CCAATGCGCTTGTATGATTTGATGACAGAGACTGTGAAGAAATCGTGAGGCGCAA 324
Db      119 ProAspAlaPheAspThrGlnGlnLeuLeuHisCysAlaGlnThrLeuLysSerGlyGln 138
Qy      325 ACGGTGAGGTGCGGACCTATGATTTGTGACACACTCAAGTTTACAGAGAGAC---ACG 381
Db      139 ProTyrGlnValProIleTyrAspPheLysThrIleGlnArgValSerAspThrPheArg 158
Qy      382 GTGGTCTACCTCGGAGCGTGTCTGTGTGAGGCACTTGTGTCTTACAGCCAGAG 441
Db      159 GlnValAlaAspAlaSerAspValIleIleLeuGlnGlyIleLeuValPheHisAspSerArg 178
Qy      442 ATCCGGACATGTTCCACTCGCGCTCTTCTGTGACACCGACCTCCGAGCTGAGCTGTCT 501
Db      179 ValArgAsnLeuMetAsnMetLysIlePheValAspThrAspAlaAspValArgLeuAla 198
Qy      502 CGAAGAGTTCTCCGGGAC---GTGCGCGGAGGAGGAGGAGCTGAGAGCATTTCTGACGCG 558
Db      199 ArgArgIleAlaGlyAspThrValGlnArgGlyArgAspValAlaAsnSerValLeuGln 218
Qy      559 TACAACACCTTGTGAAGCGCGCTTGTGAGGAGTGTGCTGCGGACAAAGAGTATGCC 618
Db      219 TyrAlaLysPheValLysProAlaPheAspPheValLeuProSerLysTyrAla 238
Qy      619 GATGTGATCATCCACGAGAGTGAGCAATGTTGTGCGATCAACCTGATGTGTGACAC 678
Db      239 AspValIleIleProArgGlyGlyAspAsnHisValAlaValAspLeuIleThrGlnHis 258
Qy      679 ATCCAGGACATTTCTG---AATGTGACATCTGCAAAATGTCACGAGAGGGTCCATGG 735
Db      259 IleHisThrLysLeuGlnGlnHisAspLeuCysLysIleTyr-----ProAsnVal 275
Qy      736 CGAGACTACAGCGGACCTTTTCTGAGCCAGGAGGACAC 774
Db      276 TyrValIleGlnSerThrPheGlnIleArgGlyMetHis 288

RESULT 5
ID      09LTY6      PRELIMINARY;      PRT;      465 AA.
AC      09LTY6;
DT      01-OCT-2000 (TREMBLrel. 15, Created)
DT      01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT      01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE      Uridine kinase-like protein.
OS      Arabidopsis thaliana (Mouse-ear cress).
OC      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC      Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC      eurosoids II; Brassicales; Brassicaceae; Arabidopsids.
OX      NCBI_TaxID=3702;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=Columbia;
RA      Sato S., Nakamura Y., Kaneko T., Kato T., Asamizu E., Tabata S.;
RL      Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
RN      [2]
RP      SEQUENCE FROM N.A.
RC      STRAIN=Columbia;
RA      MEDLINE=20277480; PubMed=10819329;
RA      Nakamura Y.;

```

```

RT      "Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence
RT      features of the regions of 4,504,864 bp covered by sixty P1 and TAC
RT      clones."
RL      DNA Res. 7:131-135 (2000).
DR      EMBL; AB024028; BA95720.1; -.
DR      HSSP; Q2698; 1BD3.
DR      InterPro; IPR006083; PRK URK.
DR      InterPro; IPR000764; Uridine_kin.
DR      Pfam; PF00485; PRK.1
DR      PRINTS; PR00988; URIDINKINASE.
DR      TIGRFAMs; TIGR00235; udk; 1.
KW      Kinase.
SQ      SEQUENCE 465 AA; 52125 MW; 4CD82DB9059E7168 CRC64;

Alignment Scores:
Pred. No.:      8.3e-41      Length:      465
Score:          568.00      Matches:      115
Percent Similarity: 66.94%      Conservative: 51
Best Local Similarity: 46.37%      Mismatches: 68
Query Match:    37.52%      Indels:      14
DB:             10      Gaps:        5

US-09-896-522-3 (1-834) x 09LTY6 (1-465)
Qy      37 GCGCGGAGGCGGACCGCTCGGACAGCGGCGCTTCTGATAGGGTGAAGCGGCGACT 96
Db      19 SerProSerAlaProAlaProLeuLysGlnProPheValIleGlyValAlaGlyGlyThr 38
Qy      97 GCCAGCGGAGAGTGCACCGTGTGTGAAGAAATCATGAGTGTCTGGACAGAAAGAGTGTG 156
Db      39 AlaSerGlyLysThrThrValCysAsnMetIleMetSerGlnLeu----- 53
Qy      157 GAACAGCGGAGGAGGAGAGTGTATCTCGAGCGAGACAGAGTTCACAAAGTCTGACG 216
Db      54 -----HisAspIleArgValAlaLeuValAlaGlnAspSerPheTyrHisSerLeuThr 71
Qy      217 GCAGACGAGAGCGGACGAGCGCTTGAAGAGCACTGATCAATTTTGACCATCCAGTGCCTTT 276
Db      72 LysGlnLysLeuAsnLysVal-----HisGlyTyrAsnPheAspHisProAspAlaPhe 89
Qy      277 GATATGATTTGATGACAGAGACTGTGAAGAAATCATGAGGCGCAAAACGTTGAGGTG 336
Db      90 AsnThrGlnValLeuLeuSerCysMetGlnLysLeuArgSerGlyIleProValAlaAsnIle 109
Qy      337 CCGACCTATGATTTTGTGACACACTCAAGTTTCCAGAGACGAGCGGTGATACCTGGG 396
Db      110 ProSerTyrAspPheLysIleHisGlnSerIleGlnSerSerSerProValAsnProGly 129
Qy      397 GACGTGTCTGTGTTGAGGCGATCTTGTGTTCTACAGCCAGAGATCCGCGGACATGTTG 456
Db      130 AspValIleIleLeuGlnGlyIleLeuValLeuAsnAspProArgValArgAspLeuMet 149
Qy      457 CACCTGGCGCTCTTGTGACACCGACTCCGAGCTGAGGCTGTCTGAAAGTTCTCCGG 516
Db      150 AsnMetLysIlePheValAspThrAspAlaAspValArgLeuSerArgIleGlnArg 169
Qy      517 GAC---GTGCGCGGAGGAGGAGCTGAGAGCACTTGAACGAGTCAACCAACCTCGTG 573
Db      170 AspThrValGlnArgGlyArgAsnIleGlnAsnValLeuGlnIleGlnThrLysPheVal 189
Qy      574 AACCGGCGCTTGAAGAGTCTGCTCGGACAAAGATGATCCGATGATGATCAACCA 633
Db      190 LysProSerPheAspLysTyrIleGlnProSerMetLysTyrAlaAspIleIlePro 209
Qy      634 CGAGAGTGAACAATATGTTGTCATCAACTGATGTGTGACAGACATCCAG---GACATT 690
Db      210 ArgGlyGlyAspAsnAspValAlaIleAspLeuIleValGlnHisIleArgThrLysLeu 229
Qy      691 CTGAATGTGACATCTGCAAAATGACACCGAGAGGTTCCAAATGGGCGGAGCTCAAGCGG 750
Db      230 CysGlnHisAsnLeuCysLysIleTyr-----SerAsnIlePheIleIleSerSer 246
Qy      751 ACCTTTTCTGAGCCAGGAGGACAC 774

```

Db 247 Threheglinlelysglymethis 254

RESULT 6

Q9LK34 PRELIMINARY; PRT; 483 AA.

AC Q9LK34; (1-OCT-2000 (TREMblrel. 15, last sequence update))

DT 01-OCT-2000 (TREMblrel. 15, last sequence update)

DT 01-MAR-2003 (TREMblrel. 23, last annotation update)

DE Uridine kinase-like protein.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.

OX NCBI_TaxID=3702;

RM [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Columbia;

RA Kaneo T., Kato T., Sato S., Nakamura Y., Asamizu E., Tabata S.;

RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=Columbia;

RX MEDLINE=20363099; PubMed=10907853;

RA Nakamura Y.;

RT "Structural analysis of Arabidopsis thaliana chromosome 3. II. RT Sequence features of the regions of 4,251,695 bp covered by ninety PL, RT TAC and BAC clones."

RL DNA Res. 7:217-221(2000).

DR EMBL; AF000381; BAB02114.1; -.

DR HSSP; Q26998; 1BD3.

DR InterPro; IPR006083; PRK_UK.

DR InterPro; IPR000764; Uridine_kin.

DR Pfam; PF00485; PRK_1.

DR PRINTS; PRO0988; URIDINKINASE.

DR TIGRFAMs; TIGR00235; udk; 1.

KM Kinase.

SQ SEQUENCE 483 AA; 54210 MW; 286277AED187020F CRC64;

Alignment Scores:

Pred. No.:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
2,29e-40	483	563.00	117	67.08%	46	16
Percent Similarity:				48.15%		
Best Local Similarity:				37.19%		
Query Match:					7	

US-09-896-522-3 (1-834) x Q9LK34 (1-483)

QY 55 CCGCAGCAGCGGCTTCTGATGAGCGGCGGCACTGCGAGGGAAGTGCACC 114

DB 59 PCHHSGIN--ProhevalilleglyvalthcgllythkrlaserglysthrThr 77

QY 115 GTGTGTGAGAAATCATGAGGTTCTGAGCAGAAAGAGTGGAAAGCGGCGGAG 114

DB 78 ValCyehasprctillelleglnlne-----Hisasphsary 90

QY 175 GTGTGATCTCTGAGCAGGAGGTTCTGAGGCTCTGAGCAGGAGGAGGAGGAG 224

DB 91 ILeValleuValaenglnasrserpethyrglyglnleuhrsrglnleuGlnhis 110

QY 235 GCCTTGAAGAGACATCAATTTTGCACATCCAGATCCCTTTGATTAATGATTGATGCAC 234

DB 111 Val-----Glnlutyraapnheasphsroaspralpheasprthnglnleu 128

QY 295 AGGACTGTGAAGAATCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 354

DB 129 HicysvalasprleleuylserserglyglnProtyglnleleProlelyrasprleu 148

QY 355 ACACACTCAAGTTACAGAGAC--ACGAGTGTACCTCGGAGAGCTGTCTGTTT 411

DB 149 ThrHicglnlrglylsValaspralphearglnValasmlacyasprVallelleleu 168

QY 412 GAGGCACTTGTGTCTTCAACAGCAGAGATCCGGAGATGTTCCAGCTGCGCTTCT 471

DB 169 GlnGlylleuValleuValhehlsasprArvalArgAspleuMetasmetlysllepe 188

QY 472 GTGACACCGACTCCGAGCTCAGCTGTCTCGAAGAGTTCTCCGGAGC--GTGCGCGCA 528

DB 189 ValAsprThraspralaspvalArgleuAlaArgArgIleArgArgasprhValGlnArg 208

QY 529 GGGAGGAGCTTGAGAGATTTCTGACGATACACCACTTCTGTGAAGCGGCTTCGAG 588

DB 209 GlyArgAsprValAsprValleuGlnGlnArgValArgleuValysProAlaPheAsp 228

QY 589 GAGTTTCTGCGGCAAGAAAGTATGCGGATGATATATCCAGAGAGAGTGCACAT 648

DB 229 AspPheValleuProserlystlyAlaAsprAlilleProalrglylAsprAsn 248

QY 649 ATGTTGCATCACTGATTCGTGAGCAGCATTCAGACATTCG--AATGGTGCATC 705

DB 249 HicValAlaValasprleuValGlnHisIleHicThlyleuGlnGlnHisAsprleu 268

QY 706 TGCAAAATGCGACCGAGAGGTTCAATGGAGCTACAGCGGACCTTTTCTGAGCCA 765

DB 269 Cyllyllelyr-----ProasprValPheValilleGlnThrPheGlnlelyArg 285

QY 766 GGGGAGCACC 774

DB 286 GlyMethis 288

RESULT 7

Q19583 PRELIMINARY; PRT; 515 AA.

AC Q19583; (1-NOV-1996 (TREMblrel. 01, Created))

DT 01-NOV-1996 (TREMblrel. 01, last sequence update)

DT 01-MAR-2003 (TREMblrel. 23, last annotation update)

DE F19B6.1a protein.

GN F19B6.1 OR F19B6.1A.

OS Caenorhabditis elegans.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;

OC Rhabditidae; Pelodertinae; Caenorhabditis.

OX NCBI_TaxID=6239;

RP SEQUENCE FROM N.A.

RA Thomas K.;

RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=99069613; PubMed=9851916;

RA none;

RT "Genome sequence of the nematode C. elegans: A platform for RT Investigating biology."

RL Science 282:2012-2018(1998).

DR EMBL; Z69635; CAA93459.1; -.

DR HSSP; Q26998; 1BD3.

DR WormRep; F19B6.1a; CB05666.

DR InterPro; IPR006082; PRK.

DR InterPro; IPR006083; PRK_UK.

DR Pfam; PF00485; PRK_1.

DR PRINTS; PRO0478; PHRIUKINASE.

DR PRINTS; PRO0988; URIDINKINASE.

DR TIGRFAMs; TIGR00235; udk; 1.

SQ SEQUENCE 515 AA; 58452 MW; 9584D947A2D3B5C5 CRC64;

Alignment Scores:

Pred. No.:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
5,19e-40	515	559.00	110	74.04%	46	2
Percent Similarity:				52.88%		
Best Local Similarity:				36.92%		
Query Match:					8	

US-09-896-522-3 (1-834) x Q19583 (1-515)

QY 61 CAGGGGCGCTTCGTATAGGGGTGAGCGCGGCACTGCGGAGGAAGTGCACCGTGTGT 120
 Db 62 LysHisProPheValIleGlyValCysGlyGlySerAlaSerGlyLysThrValAla 81
 QY 121 GAGAGATCATGAGATTGCTGGGACAGAACAGAGGTGAAACGCGGACGCGGAGGTGTC 180
 Db 82 GluLysIleValGluArgLeuGly-----IleProTrpValThr 94
 QY 181 ATCTGAGCCAGAGAGAGTCTTCAAGGTCTTGACGCGAGACGAAAGCCAGGCTTG 240
 Db 95 IleLeuSerMetAlaSerPheTyrLysValLeuThrProGluGluIleLysAlaAlaHis 114
 QY 241 AAAGACAGTCAATTTTGAACATCCAGATCCCTTTGATATGATTTGATTCACAGACT 300
 Db 115 GluSerArgTyrAsnPheAspGlyProAsnAlaPheAspPheAspLeuLysGluVal 134
 QY 301 CTGAGAACATCTGAGAGGCGAAACGAGTGGAGGTGCGGACCTATGATTTGTGACACAC 360
 Db 135 LeuLysArgLeuArgGluGlyLysSerValAspValProValTyrAspPheAsnThrHis 154
 QY 361 TCAAGGTACCAAGAGACCAAGCGGTGTCTACCTCGGACGTTGTTCTGTTGAGGCAATC 420
 Db 155 SerArgAspProAsnSerLysMetMetTyrGlyAlaAspValLeuIlePheGluGlyIle 174
 QY 421 TTGATGTTCTACAGCAGAGAGATCCGAGACATGTTCCACCTGCGCTTGTGTGACACCC 480
 Db 175 LeuAlaPheHisAspGluArgIleLysAsnLeuMetAspMetLysValPheValAspThr 194
 QY 481 GACTCCGACGTCAAGCTGTCTCGAAGATTCTCCGGACGTCGCGC---CGAGGAGGAGAC 537
 Db 195 AspGlyAspLeuArgLeuAlaArgIleValArgAspValThrAspArgIleArgAsp 214
 QY 538 CTGAGACAGATTCTGACGACATCACACCTTCTGAAAGCCGCGCTTGCAGAGACTTGC 597
 Db 215 IleAspGlyIleMetGluGlnTyrPheThrPheValLysProAlaPheAspLysTyrIle 234
 QY 598 CTGCGCAAAAGAGATGTCGATGTCATCCACGAGAGAGTGGCAAAATATGTTGCC 657
 Db 235 AlaProCysMetAspSerAlaAspLeuIleValProArgIleGlyLysAsnAspValAla 254
 QY 658 ATCAACCTGATCTGCGACACATC 681
 Db 255 IleAspMetIleValGlnAsnVal 262

RESULT 8

QY09317 PRELIMINARY; PRT; 555 AA.
 ID Q9317 AC Q9317;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE F19B6.1b protein.
 GN F19B6.1 OR F19B6.1B.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
 OC Rhabditidae; Peloderiinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RN SEQUENCE FROM N.A.
 RA Thomas K.;
 RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99069613; PubMed=9851916;
 RA none;
 RT "Genome sequence of the nematode C. elegans: A platform for
 RT investigating biology.";
 RT Science 282:2013-2018 (1998).
 DR EMBL; Z69635; CA93462.1; -.
 DR HSSP; Q26998; 1BD3.
 DR WormPep; F19B6.1b; CE20700.

DR InterPro; IPR006082; PRK.
 DR InterPro; IPR006083; PRK URK.
 DR InterPro; IPR000764; Uridine_kin.
 DR Pfam; PF00485; PRK.1.
 DR PRINTS; PR00478; PHRIDKINASE.
 DR PRINTS; PR00988; URIDINKINASE.
 DR TIGRFAMs; TIGR00235; udk; 1.
 DR SEQUENCE 555 AA; 62673 MW; D0786FAB98B8CF98 CRC64;

Alignment Scores:

Pred. No.:	Score:	Percent Similarity:	Best Local Similarity:	Query Match:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
5.29e-40	559.00	74.04%	52.88%	36.92%	555	110	44	46	8	2

US-09-896-522-3 (1-834) x Q9317 (1-555)

QY 61 CAGGGGCGCTTCGTATAGGGGTGAGCGCGGCACTGCGGAGGAAGTGCACCGTGTGT 120
 Db 102 LysHisProPheValIleGlyValCysGlyGlySerAlaSerGlyLysThrValAla 121
 QY 121 GAGAGATCATGAGATTGCTGGGACAGAACAGAGGTGAAACGCGGACGCGGAGGTGTC 180
 Db 122 GluLysIleValGluArgLeuGly-----IleProTrpValThr 134
 QY 181 ATCTGAGCCAGAGAGAGTCTTCAAGGTCTTGACGCGAGACGAAAGCCAGGCTTG 240
 Db 135 IleLeuSerMetAlaSerPheTyrLysValLeuThrProGluGluIleLysAlaAlaHis 154
 QY 241 AAAGACAGTCAATTTTGAACATCCAGATCCCTTTGATATGATTTGATTCACAGACT 300
 Db 155 GluSerArgTyrAsnPheAspGlyProAsnAlaPheAspPheAspLeuLysGluVal 174
 QY 301 CTGAGACAGATTCTGACGACATCACACCTTCTGAAAGCCGCGCTTGCAGAGACTTGC 597
 Db 175 LeuLysArgLeuArgGluGlyLysSerValAspValProValTyrAspPheAsnThrHis 194
 QY 361 TCAAGGTACCAAGAGACCAAGCGGTGTCTACCTCGGACGTTGTTCTGTTGAGGCAATC 420
 Db 195 SerArgAspProAsnSerLysMetMetTyrGlyAlaAspValLeuIlePheGluGlyIle 214
 QY 421 TTGATGTTCTACAGCAGAGAGATCCGAGACATGTTCCACCTGCGCTTGTGTGACACCC 480
 Db 215 LeuAlaPheHisAspGluArgIleLysAsnLeuMetAspMetLysValPheValAspThr 234
 QY 481 GACTCCGACGTCAAGCTGTCTCGAAGATTCTCCGGACGTCGCGC---CGAGGAGGAGAC 537
 Db 235 AspGlyAspLeuArgLeuAlaArgIleValArgAspValThrAspArgIleArgAsp 254
 QY 538 CTGAGACAGATTCTGACGACATCACACCTTCTGAAAGCCGCGCTTGCAGAGACTTGC 597
 Db 255 IleAspGlyIleMetGluGlnTyrPheThrPheValLysProAlaPheAspLysTyrIle 274
 QY 598 CTGCGCAAAAGAGATGTCGATGTCATCCACGAGAGAGTGGCAAAATATGTTGCC 657
 Db 275 AlaProCysMetAspSerAlaAspLeuIleValProArgIleGlyLysAsnAspValAla 294
 QY 658 ATCAACCTGATCTGCGACACATC 681
 Db 295 IleAspMetIleValGlnAsnVal 302

RESULT 9

QY08VB2 PRELIMINARY; PRT; 466 AA.
 ID Q8VB2 AC Q8VB2;
 DT 01-MAR-2002 (TREMBlrel. 20, Created)
 DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE Putative uracil phosphoribosyltransferase.
 GN ATIG55810.
 OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC Euroside II; Brassicales; Brassicaceae; Arabidopsids.
 OC NCBI_TaxId=3702;
 RN
 RP SEQUENCE FROM N.A.
 RA Yamada K., Banh J., Chan M.M., Chang C.H., Chang E., Dale J.M.,
 RA Dang J.M., Goldsmith A.D., Lee J.M., Onodera C.S., Quach H.L.,
 RA Tang C., Toriumi M., Wu H.C., Yamamura Y., Yu G., Bowser L.,
 RA Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
 RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
 RA Meyers T.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,
 RA Sakurai T., Satou M., Seki M., Shim P., Southwick A., Shinzaki K.,
 RA Davis R.W., Ecker J.R., Theologis A.;
 RT "Full Length cDNA of gene At1g55810 (GT:1522278).";
 RT Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 RL
 RN
 RP SEQUENCE FROM N.A.
 RA Yamada K., Chan M.M., Chang C.H., Dale J.M., Deng J.M., Hsuan V.W.,
 RA Lee J.M., Quach H.L., Tang C., Toriumi M., Wu H.C., Yu G., Bowser L.,
 RA Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
 RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
 RA Miranda M., Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M.,
 RA Seki M., Shim P., Southwick A., Shinzaki K., Davis R.W., Ecker J.R.,
 RA Theologis A.;
 RT "Arabidopsis Open Reading Frame (ORF) Clones";
 RT Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
 RL
 RN
 RP EMBL; AY072218; AL60039.1; -;
 DR EMBL; AY122946; AAM67479.1; -;
 DR InterPro; IPR006083; PRK URK.
 DR InterPro; IPR00764; Uridine_kin.
 DR Pfam; PF00485; PRK; 1.
 DR PRINTS; PRO0988; URIDINKINASE.
 DR TIGRFAMs; TIGR00235; udk; 1.
 KW Glycosyltransferase; Transferase.
 SQ SEQUENCE 466 AA; 52443 MW; 3FA6783CB49727P7 CRC64;
 Alignment Scores:
 Pred. No.: 1,02e-38 Length: 466
 Score: 544.00 Matches: 116
 Percent Similarity: 63.16% Conservative: 52
 Best Local Similarity: 43.61% Mismatches: 78
 Query Match: 35.93% Indels: 20
 Gaps: 8
 DB: NCBI_TaxId=7227;
 US-09-896-522-3 (1-834) x Q8VYB2 (1-466)
 QY 31 AGCCCCGCGCGAG-----GCCGACCGTCGACGCGCGCCCTTCTGATGAGG 81
 Db 29 SerAsnAspProGluGlnMetAlaGluGlnGlnHisGlyGlnProPheValIleGly 48
 QY 82 GTGAGCGCGGCGACTGCCACCGGGAAGTGCACCGTGTGTGAAGAATCATGAGTGTCTG 141
 Db 49 ValAlaGlyValAlaAspSerGlyThrThrValCysAspMetIleMetGlnGlnLeu 68
 QY 142 GGAACAAGACAGGTGGAACAAGCGCGGAGAGGTGTCATCTCGAGCCGACGAGTTC 201
 Db 69 -----HisAspGlnAspAlaValAlaValAlaGlnAspSerPhe 81
 QY 202 TACAAGGTCCTGACGAGGAGCAAGAGCGCAAGCCTTGAAGAAGACATCAATTTTAC 261
 Db 82 TyHHisAsnValAsnGluValGluLeuValAlaValAla-----HisAspTyrAsnAsp 99
 QY 262 CATCCGATGCGCTTTGATTAATGATTGATGACACGAGCTGTGAAGAATCATGAGAGGC 321
 Db 100 HisProAspAlaPheAspThrGlnGlnLeuLeuSerSerMetGluLysLeuArgGly 119
 QY 322 AAAACGCTGAGGAGCGACCTATGATTGTGACACACCTCAAG-----TACACGAG 375
 Db 120 GlnAlaValAspIleProAsnTyrAspPheLysSerTyrLysAsnValAlaPheProPro 139
 QY 376 ACCACGGTGTCTTACCTGCGAGCGTGGTCTGTGAGGAGCATCTTGTTCTACAGC 435

Db 140 ArgArg---ValAsnProSerAspValIleIleLeuGluGlyIleLeuIlePheHisAsp 158
 QY 436 CAGGAATCCGGGACACATGTTCCACCTCGCCCTTGTGACACGACCTCCAGCGCAAG 495
 Db 159 ProArgValArgAspLeuMetAsnMetLysIlePheValAlaAspAlaAspValArg 178
 QY 496 CTGTCTCGAAGAGTTCTCCGGGAC---GTGCGCGGAGGAGGAGGAGGAGCATTTCTG 552
 Db 179 LeuAlaArgArgIleLysArgAspThrValGluGlyGlyArgAspIleAlaThrValLeu 198
 QY 553 ACCGACTACACCACTTGTGAAAGCGGCTTGTGAGAGTTCCTGCGCAAGAAAG 612
 Db 199 AspGlnTyrSerLysPheValLysProAlaPheGluAspPheIleLeuProThrLysLys 218
 QY 613 TATGCCGATGTATCATCCACGAGAGTGACAAATATGTTGCCATCAACCTGATCTG 672
 Db 219 TyrAlaAspIleIleIleProArgGlyGlyAspAsnHisValAlaIleAspLeuIleVal 238
 QY 673 CACGACATCCAGGACATTTCTG--AATGGTACATCTGCAAAATGCGAGGAGGATCC 729
 Db 239 GlnHisIleHisThrLysLeuGlyGlnHisAspLeuGlyAlaIleTyr-----Pro 255
 QY 730 AATGGCGGAGCTACAAAGCGACCTTTTGTGAGCGGAGGAGCACCCTGGATCTGACC 789
 Db 256 AsnLeuTyrValIleGlnSerThrPheGlnIleArgGlyMetHisThrLeuIleArgAsp 275
 QY 790 TCTGGCAACCGTCACAT 807
 Db 276 SerLysThrThrLysHis 281
 RESULT 10
 ID Q9V810 PRELIMINARY; PRT; 614 AA.
 AC Q9V810; Q9V811;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE CG4798 protein.
 GN CG4798.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OC NCBI_TaxId=7227;
 RN
 RP SEQUENCE FROM N.A. (LONG AND SHORT ISOFORMS).
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktoglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
 RA Burks K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport I.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster G., Gabrielian A.B., Gary N.S., Galbati W.M., Glasser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Houlton D., Houston K.A., Howland T.J., Wei M.-H., Ibeagham C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lascko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,

RA Nelson R., Nelson K.A., Nixon K., Nuskern D.R., Paclzel J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Pui V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Stimpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stepleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner A., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveli J.S., Zhan M., Zhang G., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
RT "The genome sequence of *Drosophila melanogaster*.";
RL Science 287:2185-2195(2000).
CC -1. ALTERNATIVE PRODUCTS: 2 ISOFORMS, A LONG FORM (SHOWN HERE) AND A
CC SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC EMBL; AE003803; AAF57871.1; -;
DR EMBL; AE003803; AAF57873.1; -;
DR HSSP; Q26998; 1BD3.
DR FlyBase; FBGm0034213; CG4798.
DR InterPro; IPR006083; PRK_URK.
DR InterPro; IPR00764; Uridine_kin.
DR Pfam; PF00485; PRK; 1.
DR PRINTS; PR00988; URIDINKINASE.
DR TRIFAPMS; TIGR00235; udk; 1.
KW Alternative splicing, Hypothetical protein.
FT VASPLIC 1 207 MISSING (IN SHORT ISOFORM).
SQ SEQUENCE 614 AA; 68704 MW; C81B63B9DBE332A1 CRC64;

Alignment Scores:	
Pred. No.:	1,65e-38
Score:	342.00
Percent Similarity:	71.84%
Best Local Similarity:	51.46%
Query Match:	35.80%
DB:	5
	Gaps: 3
	Length: 614
	Matches: 106
	Conservative: 42
	Mismatches: 50
	Indels: 8
	Gaps: 3

US-09-896-522-3 (1-834) x Q9V810 (1-614)

QY 67 CCCCTCCTGATAGGGGTGAGCGGGCGGCACCTGCCAGCCGGGAAGTCGACCCTGTGTGTGAGAAG 126
|||:::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 174 ProPheValIleGlyIleCysGlyGlySerAlaSerGlyLysThrTrnValAlaGluLys 193

```

QY      127 ATCATGAGTGTCTGGACACAGAGTGGAAACGCGGACGGAGAGTGTCTCATCTG 186
          |||::||| ||| ::||| ||| ::|||
Db      194 ILeIleGluSerLeu-----AapValProTrp-----ValThrLeuLeu 206

```

Db
oy 187 AGCCAGGACAGGTTTCACAAAGGTCCTGACCGCAGAGCAGAGAGAGGCGCAAGGCGTTGAAGA 246
||| ||| |||||||||:::||||| ::::|
207 SerMetAspCysPheTyrLysIleLeuAsnGluLysGlnHisGluGlnAlaLeuIleAsn 226

24 CAGTACCAATTTTGACCAATCCAGATGCTTGGATATATGTTTGGATGCAAGAGATCTGGAG 306
:::|||||:::|||||
227 GluTyrAsnHisAspHisProAspAlaPheAspIleGluLeuLeuLeuAspValLeuThr 246

Dd 247 LysLeuLysGluGIYArgLysValGluValProValTyrAsnPhenValThrHisGIYArg 266

Db 267 G L U S E R G I n t h r L y s t h r M e t y r G l y A l a s n V a l l e I e p h e g i n g I y l e l e u t h r 286

Db 287 PheHisSerProGluValLeuLysLeuLysAspMetLysIlePheValAspThrAspPro 306

[illegible]

Db 327 GlyValLeuysGlnTyrLeuAsnMetValLysProSerTyrCysAsnTyrIleAlaPro 346

[illegible]

RESULT 11
Q8MOK4
ID Q8MOK4
PRELIMINARY;
PRT; 626 AA.

DT	01-OCT-2002	(TREMBLrel. 22, Created)
DT	01-OCT-2002	(TREMBLrel. 22, Last sequence update)
DT	01-MAR-2003	(TREMBLrel. 23, Last annotation update)
DE	10035950	

GN CG4/98.
 OS *Drosophila melanogaster* (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 0C Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

CC 2pnyatvttca, ctggcgaatttacc, ctggcgaatttacc;
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.

RA Seaplecon M., Broksstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,

RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celinker S.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DOI: 10.1093/nar/gkf770

DR FlyBase:FBGN0034213; CG4798.
DR InterPro:IPR006083; PRK_UK.
DR InterPro:IPR00764; Uridine_kin.
DR Pfam:PF00485; PRK_1

DR FRINIS; FR00268; URIDININAME,
DR TIGRFAMS; TIGR00235; udk; 1.
SQ SEQUENCE 626 AA; 70001 MW; 9D8EC0A7331A6EF3 CRC64;

Pred. No.:	1.65e-38	Length:	626
Score:	542.00	Matches:	106
Percent Similarity:	71.84%	Conservative:	42

Query Match:	35.80%	Indels:	8
DB:	5	Gaps:	3

67 CCCTTCGTAGGGGTGAGCGGCGCACTGCCAGCGGAAGTGCACCGTGTGAGAG 126

Qy 127 ATCATGAGTTCCTGGGACAGACGAGTGGAACAGCGGCAGCGAAGCTGGTATCCTG 186
|||::||| ||| ::||| ||| ::|||
Nb 206 TTTCTAGTGSAAV-----ASNVAPRDTY-----VAITHTRLNEN 218

QY 187 AGCCAGACAGTTCCTACAGGCTCTGACGGCAGACAGAGGCCAAGCCTTGAAAGA 246
||| ||| |||||:::| |::| |::| |||
Db 219 SerMetAsnCvSpHetYrvsIleLeuSngIuYvsgIhIsgIuGInAlaLeuIeAsn 238

QY	247	CAGTACAA	TTTTGCA	CCATCAGA	GCCCTTGAT	AATGATT	TGTATGCAC	AGACTCTGA	AG	306
	:	:	:	:	:	:	:	:	:	:
Db	239	GLuTYrAsn	PheAsn	PHisPro	AlaPhe	AspIle	GLuleuLeu	eudArVal	IeuThr	258

Dy 307 AACATCGTGGAGGCGAACAACGGTGAGGCGCCGACTATTCCTTTGTGAACACACTCAAGG 366
::: |||||::: ||||| ||||| ||||| |||||
Db 259 LysleuylsgluArglyValgluValProvalTyAsnPhelThrhisclArg 278

367 TTACGAGACCAAGGTGCTCTACCCCTGGGACGTTGTTCTGTTGAGGCATCTTGGTG 426

```

Db      279 GluSerGlnThrLysThrMetTyrGlyAlaSerValIleLeuPheGlnGlyIleLeuThr 298
      427 TTCTACAGCCGAGAGATCCGGACATGTCACCTGGCCCTTTCTGGACACCGACTCC 486
      299 PheHisSerProGlnValLeuLysLeuLeuAspMetLysIlePheValAspThrAspPro 318
      487 GAGTCAGGCTGTCGTGACAGAGATTCTCGGGAGCTG---GGCGAGGAGGAGGAGCTGGAG 543
      319 AspIleArgLeuAlaArgAlaArgLeuArgAspIleSerGlnArgIleValArgAspLeuLys 338
      544 CAGATTCTGACGAGATACACACCTTGAGAGCCGAGCTTCGAGAGATTCTGCTGCCG 603
      339 GlyValLeuLysGlnTyrLeuAsnMetValLysProSerTyrCysAsnTyrIleAlaPro 358
      604 ACAAGAAGATGATCCGATGTCATATCCACGAGAGAGTGAACAATATGTTCCCATCAAC 663
      359 ThrMetAlaHisAlaAspIleIleValProArgGlyGlyAspAsnLysValAlaIleHis 378
      664 CTGATGTCGACGACATC 681
      379 LeuIleValGlnHisVal 384

```

RESULT 12

```

Q8N524  PRELIMINARY;  PRT;  419 AA.
AC  Q8N524
DT  01-OCT-2002 (TrEMBLrel. 22, Created)
DT  01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DE  01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE  Similar to uridine kinase-like 1.
OS  Homo sapiens (Human).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX  NCBI_TaxID=9606;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  TISSUE=Placenta;
RA  Straussberg R.;
RL  Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR  EMBL; BC03078; AAH3078.1; -.
DR  InterPro; IPR006083; PRK_UK.
DR  InterPro; IPR000764; Uridine_kin.
DR  Pfam; PF00485; PRK_1.
DR  PRINTS; PR00478; PHRIDUKINASE.
DR  PRINTS; PR00988; URIDINKINASE.
DR  TIGRPFAM; TIGR00235; udk; 1.
KM  Kinase.
SQ  SEQUENCE 419 AA; 45972 MW; 1D1F1E75305AD25 CRC64;

```

Alignment Scores:

```

Pred. No.:      2e-38      Length:      419
Score:          540.50     Matches:      110
Percent Similarity: 64.20%  Conservative: 46
Best Local Similarity: 45.27%  Mismatches: 60
Query Match:    35.70%     Indels:      27
DB:             4         Gaps:      5

```

US-09-896-522-3 (1-834) x Q8N524 (1-419)

```

QY  25 TGGGAGAGCCCGCCCGCG-----GAGGCCGACCGT 54
      |||:|||||
Db  65 CysLysSerGlnProPheLeuLysArgThrSerLysArgThrIleTyrThrAlaGlyArg 84
      |||:|||||
QY  55 CCG-----CACCAGCGCCCTTCTCGATAGGGGTGAGC 87
      |||:|||||
Db  85 ProProTyrThrAsnGlnHisGlyThrGlnSerLysLeuAlaPheAlaIleGlyLeuGly 104
      |||:|||||
QY  88 GGGGGAAGTCCCGAGCGGAGAGTGCACCGGTGTGAGAAATCATGAGTTGCTGGAGCAG 147
      |||:|||||
Db  105 GlyGlySerLysSerGlyLysThrThrValAlaArgMetIleIleGlnAlaLeu----- 122

```

```

QY  148 AACGAGTGAACAAGCGGAGCGGAGAGGTGTCATCTGAGCCAGGAGTTCTACAG 207
      :|||
Db  123 ---AspValProTrp-----ValValLeuLeuSerMetLysPheTyrLys 137
      |||:|||||
QY  208 GTCTTCAGCGGAGAGAGAGGCGGAGGCGCTTGAAGAAGACATGACATTTTGAACATCA 267
      |||:|||||
Db  136 ValLeuThrGlnGlnGlnGlnGlnGlnAlaIleHisAsnHisPheAsnHisPheAspPro 157
      |||:|||||
QY  268 GATGCTTTGATATATATTTGATGACACAGACTCTGAACATCTGAGAGGCGGAGGAG 327
      |||:|||||
Db  158 AspAlaPheAspPheAspLeuIleIleSerThrLeuLysLeuLysGlnGlyLysSer 177
      |||:|||||
QY  328 GTGAGAGTGGCGGACCTATGATTTTGTGACACACTCAAGGTTACGAGACCAAGGTGTC 387
      |||:|||||
Db  178 ValLysValProLefTyrAspPheThrThrHisSerArgLysLysAspTyrLysThrLeu 197
      |||:|||||
QY  388 TACCTCGGAGCGCTGTGTTGAGGAGCATTTGCTTTCACGAGAGATCCGG 447
      |||:|||||
Db  198 TyrGlyAlaSerValIleLeuPheGlnGlyIleMetAlaPheAlaAspLysThrLeuLeu 217
      |||:|||||
QY  448 GACATGTTCCACTGCGCTCTTCTGACACCGACTCCGAGCTGCTGCGAAGA 507
      |||:|||||
Db  218 GluLeuLeuAspMetLysIlePheValAspThrAspSerAspIleArgLeuValArgArg 237
      |||:|||||
QY  508 GTTCTCGGAGAGCTG---GGCGAGGAGGAGGAGCTTGAGAGATTTCTGACGACATCAC 564
      |||:|||||
Db  238 LeuArgArgAspIleSerGlnAlaArgAlaArgAspIleGlyValIleLysGlnTyrAsn 257
      |||:|||||
QY  565 ACCTTGAGAGGCGGCGCTTCGAGAGATTGCTGCTGCGAACAAGATGCGCATGTG 624
      |||:|||||
Db  258 LysPheValLysProSerPheAspGlnTyrIleGlnProThrMetArgLeuAlaAspIle 277
      |||:|||||
QY  625 ATATCTCCACGAGAGATGACAAATATGTTGCCATTAACCTGATCTGTCGACGACATCCAG 684
      |||:|||||
Db  278 ValValProArgGlySerGlyAsnThrValAlaIleAspLeuIleValGlnHisValHis 297
      |||:|||||
QY  685 GACATCTG 693
      |||
Db  298 SerGlnLeu 300

```

RESULT 13

```

Q8LD95  PRELIMINARY;  PRT;  469 AA.
AC  Q8LD95
DT  01-OCT-2002 (TrEMBLrel. 22, Created)
DT  01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DE  01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE  Putative uracil phosphoribosyl transferase.
OS  Arabidopsis thaliana (Mouse-ear cress).
OC  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC  eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX  NCBI_TaxID=3702;
RN  [1]
RP  SEQUENCE FROM N.A.
RA  Haas B.J., Volfovsky N., Town C.D., Troupkan M., Alexandrov N.,
RA  Feldman K.A., Flavell R.B., White O., Salzberg S.L.;
RT  "Full-length messenger RNA sequences greatly improve genome
RT  annotation.";
RL  Genome Biol. 0:0-0 (2002).
RN  [2]
RP  SEQUENCE FROM N.A.
RA  Brover V., Troupkan M., Alexandrov N., Lu Y.-P., Flavell R.,
RA  Feldman K.;
RL  "Full-length cDNA from Arabidopsis thaliana.";
RL  Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR  EMBL; AY086133; AAM63338.1; -.
DR  InterPro; IPR006083; PRK_UK.
DR  InterPro; IPR000764; Uridine_kin.
DR  Pfam; PF00485; PRK_1.
DR  PRINTS; PR00988; URIDINKINASE.
DR  TIGRPFAM; TIGR00235; udk; 1.
KM  Transferase.

```



```

Db      120 LeuArgysglglnAlaValAspIleProAsnIlyrAspHeuSseryTlyAsAsn 139
Qy      367 ---TTACAGAGACACGAGTGTACCTCGGAGCTGTCTGTTGAGGAGCATTTG 423
Db      140 ValPheProArgArg---ValAsnProSerAspValIleIleLeuGlnGlyIleLeu 158
Qy      424 GTGTTACAGACGAGAGATCCCGAGATGTTCCAGCTGGCCCTTCCTGGAGACC--- 480
Db      159 IlePheHisAspProArgValArgAspLeuMetMetCylsIlePheValAspAlaGly 178
Qy      480 ----- 480
Db      179 LeuSerHisThrLysProValAsnThrTyrValValYsSerValAlaTyrMetArgArg 198
Qy      481 -----GACTCCGAGCTGAGCTGTCTGAGAGATTTCTCCG 516
Db      199 CysThrCysIleCysThrHisGlnAspAlaAspValArgLeuAlaArgGlyIleArg 218
Qy      517 GAC---GTGGCGGAGGAGGAGGAGCTGAGAGATTTGACGAGTACACCACTTCGTG 573
Db      219 AspThrValGlnLysGlnArgAspIleAlaThrValLeuAspGlnTyrSerLysPheVal 238
Qy      574 AAGCGGCTTCGAGAGATTTCTGCTGCGGAGAGAGATGAGTATGCGATGATCATCCA 633
Db      239 LysPheAlaPheGlnAspPheIleLeuProThrLysTyrAlaAspIleIleIlePro 258
Qy      634 CGAGAGATGACATATGTTGCTGATCACTGATGCTGACAGCATCCAGCATTTG 693
Db      259 ArgGlyGlnAspAsnHisValAlaIleAspLeuIleValGlnHisIleHisThrLysLeu 278
Qy      694 ---AATGTCGACATCTGCAATATGCGACCGAGAGGATTCATAGGCGAGACTCAACCG 750
Db      279 GlnGlnHisAspLeuCysLysIleTyr-----ProAsnLeuTyrValIleGlnSer 295
Qy      751 ACCTTTTCTGAGCGAGGAGGAGCACTGAGATGCTGATGCTGCAACGCTGACAT 807
Db      296 ThrPheGlnIleArgGlnMetHisThrLeuIleArgAspSerLysThrThrLysHis 314

RESULT 15
Q9LFZ2 PRELIMINARY; PRT; 1060 AA.
ID Q9LFZ2;
AC Q9LFZ2;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE F20N2.19.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC Eurosid II; Brassicales; Brassicaceae; Arabidopsis.
OC NCBI_TaxID=3702;
RN
RP SEQUENCE FROM N.A.
RA Khan S., Brooks S., Buehler E., Chao Q., Johnson-Hopson C., Kim C.,
RA Shim P., Altafi H., Bei O., Chin C., Chion J., Choi E., Conn L.,
RA Conway A., Gonzalez A., Hansen N., Howng B., Koo T., Lam B., Lee J.,
RA Lanz C., Li J., Liu A., Liu K., Liu S., Mukharaky N., Nguyen M.,
RA Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaveri A.,
RA Toriumi M., Vayberg M., Yu G., Federspiel N.A., Theologis A.,
RA Ecker J.R.;
RT "Genomic sequence for Arabidopsis thaliana BAC F20N2 from chromosome
RT 1.";
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
RN
RP SEQUENCE FROM N.A.
RA Ecker J.R.;
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
RN
RP SEQUENCE FROM N.A.
RA Ecker J.R.;
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
RN

```

```

RP SEQUENCE FROM N.A.
RA Khan S., Buehler E.E., Dunn P.P., Feng J.J., Kim C.C., Li Y.Y.,
RA Shim P.P., Altafi H.H., Araujo R.R., Conn L.L., Conway A.A.B.,
RA Walker M.M., Altafi H.H., Araujo R.R., Conn L.L., Conway A.A.B.,
RA Gonzalez A.A., Hansen N.N.P., Huizar L.L., Kremetskaya I.I.,
RA Lanz C.C., Li J.J., Liu S.S., Luoro S.S., Rowley D.D., Schwartz J.J.,
RA Toriumi M.M., Vayotskaya V.V., Yu G.G., Davis R.R.W.,
RA Federspiel N.N.A., Theologis A.A., Ecker J.R.;
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
RN
RP SEQUENCE FROM N.A.
RA Ecker J.R.;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
RN
RP SEQUENCE FROM N.A.
RA Cheuk R., Shim P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C.,
RA Khan S., Kim C., Altafi H., Bei O., Chin C., Chion J., Choi E.,
RA Conn L., Conway A., Gonzalez A., Hansen N., Howng B., Koo T., Lam B.,
RA Lee J., Lanz C., Li J., Liu A., Liu K., Liu S., Mukharaky N.,
RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,
RA Thaveri A., Toriumi M., Vayberg M., Yu G., Davis R., Federspiel N.,
RA Theologis A., Ecker J.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC002328; AAF79498.1; -.
DR HSSP; Q26998; 1BD3.
DR InterPro; IPR006083; PRK_URK.
DR InterPro; IPR000764; Uridine_kin.
DR Pfam; PF00485; PRK_1.
DR PRINTS; PR00988; URIDINKINASE.
SQ SEQUENCE 1060 AA; 116160 MW; 03982556AB31D6B CRC64;

Alignment Scores:
Score: 2,61e-35 Length: 1060
Percent Similarity: 506.00 Matches: 109
Best Local Similarity: 59.18% Conservative: 49
Query Match: 40.82% Mismatches: 59
Query Match: 33.42% Indels: 50
DB: Gaps: 9

US-09-896-522-3 (1-834) x Q9LFZ2 (1-1060)
Qy      31 AGCCCGCGCGCGAG-----GCCGACCGTCCGACGAGCGCCCTTGATA--- 78
Db      29 SerAsnArgProGlnGlnMetAlaGlnGlnGlnGlnHisGlnIleProPheValIleGly 48
Qy      79 -----GGGTGACCGCGCGCGACTCCGCGGAGATCGACCGTGTGTAAGATC 129
Db      49 GluSerArgGlyValAlaGlyAlaIleAspGlyLysThrThrValCysAspMetIle 68
Qy      130 ATGAGTGTGTGGGAGACAGAAACGAGTGAACAGCGGACCGAGAGTGGTGCATCGAGC 189
Db      69 MetGlnGlnLeu-----HisAspGlnArgAlaValAlaValAsn 81
Qy      190 CAGACAGGTTTTCACAGGCTCTGACGAGGAGAGAGCCAGCTTGAAGAGACAG 249
Db      82 GlnAspSerPheThrHisAsnValaLanGlnValaLanGlnValaLanGlnValaLan 99
Qy      250 TACATTTTACCATCATCATGCTCTTATGATTTATGATGACAGAGACTTGAAGAAC 309
Db      100 TyAsnPheAspHisProAspAlaPheAspThrGlnGlnLeuSerMetGluLys 119
Qy      310 ATGTCGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 366
Db      120 LeuArgLysGlnAlaValAspIleProAsnIlyrAspHeuSseryTlyAsAsn 139
Qy      367 ---TTACAGAGACACGAGTGTACCTCGGAGCTGTCTGTTGAGGAGCATTTG 423
Db      140 ValPheProArgArg---ValaAsnProSerAspValIleIleLeuGlnGlyIleLeu 158
Qy      424 GTGTTTACAGACGAGAGATCCCGAGATGTTCCAGCTGGCCCTTCCTGGAGACC--- 480
Db      159 IlePheHisAspProArgValArgAspLeuMetMetCylsIlePheValAspAlaGly 178
Qy      480 ----- 480

```

```
Db      179 LeuSerHisThrLysProValAsnThrTyrValValLysSerValAlaTyrMetArgArg 198
QY      481 -----GACTCCGACGTCAGGCTGTCTCCGAAGAGTCTCCGG 516
Db      199 CysThrCysIleCysThrHisGluAspAlaAspValArgLeuAlaArgArgIleLysArg 218
QY      517 GAC---GTGCGCCGAGGAGGAGACCTTGAGCAGATTCTGACGACATACACCACCTTCGTG 573
Db      219 AspThrValGluLysGlyArgAspIleAlaThrValLeuAspGlnTyrSerLysPheVal 238
QY      574 AAGCCGGCCTTGAGGAGATTCTGCCTGCCGACAAAGATATGCCGATGTGATCATCCCA 633
Db      239 LysProAlaPheGluAspPheIleLeuProThrLysTyrAlaAspIleIleIlePro 258
QY      634 CGAGGAGTGCACATATGTTGCCATCAACCTGATCGTCAGACATCCAGACATTCTG 693
Db      259 ArgGlyGlyAspAsnHisValAlaIleAspLeuIleValGlnHisIleHisThrLysLeu 278
QY      694 --AATGGTGACATCTGCCAA 711
Db      279 GlyGlnHisAspLeuCysLys 285
```

Search completed: November 25, 2003, 07:39:11
Job time : 69.1806 secs


```

RT mRNA:
RU Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Phosphorylates uridine and cytidine to uridine
CC monophosphate and cytidine monophosphate. Does not phosphorylate
CC deoxyribonucleosides or purine ribonucleosides. Can use ATP or GTP
CC as a phosphate donor. Can also phosphorylate cytidine and uridine
CC nucleoside analogs such as 6-azauridine, 5-fluorouridine, 4-
CC thioridine, 5-bromouridine, N(4)-acetylcytidine, N(4)-
CC benzoylcytidine, 5-fluorocytidine, 2-thiocytidine, 5-
CC methylcytidine, and N(4)-antiseiocyridine.
CC -1- CATALYTIC ACTIVITY: ATP + uridine = ADP + UMP.
CC -1- CATALYTIC ACTIVITY: ATP + cytidine = ADP + CMP.
CC -1- PATHWAY: Pyrimidine salvage pathway.
CC -1- TISSUE SPECIFICITY: Ubiquitous.
CC -1- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF237290; AAK8334.1; -
DR EMBL: AF254133; AAK49122.1; -
DR EMBL: AK022317; BAB14010.1; -
DR EMBL: AF125106; AAL75943.1; -
DR InterPro: IPR006083; PRK_UK.
DR InterPro: IPR00764; Uridine_kin.
DR Pfam: PF00485; PRK; 1.
DR PRINTS: PR00988; URIDINKINASE.
DR TIGRfams: TIGR00235; udk; 1.
KM Transferase; Kinase; ATP-binding.
FT NP BIND 30 37 ATP (POTENTIAL).
FT CONFLICT 8 17 DESPAREAD -> GARAPARA (IN REF. 4).
FT CONFLICT 56 57 OR -> HG (IN REF. 4).
FT CONFLICT 247 247 S -> T (IN REF. 4).
SQ SEQUENCE 277 AA; 31434 MW; AFD9ED92780CD502 CMC64;

Alignment Scores:
Pred. No.: 1,1e-101 Length: 277
Score: 1450.00 Matches: 277
Percent Similarity: 100.00% Conserves: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 95.77% Indels: 0
Gaps: 0

US-09-896-522-3 (1-834) x UCK1_HUMAN (1-277)
QY 1 ATGGCTTCGGGGGAGGCGAAGATCGGAGCGCCCGCGGAGCGGACCGTCCGCAC 60
DB 1 MetlaSerlaaGlyGluAspCysGluSerProAlaProGluAlaAspArgProHis 20
QY 61 CAGGGGCGCTTCGATAGGGGGTGAAGCGGCGGACTCGGACGGGGAAGTGCAGCGTGT 120
DB 21 GlnArgProPheLeuIleGlyValSerCylGlyThrAlaSerGlyLysSerThrValCys 40
QY 121 GAGAAATCATGAGATTGCTGGACAGAACGAGGTGGAACGCGGACGCGAGAGTGTG 180
DB 41 GluLysIleMetGluLeuLeuGlyGlnAlaSerGlnValGluGlnAlaGlyValVal 60
QY 181 ATCTGAGCCGAGGACAGGTTTACAAGGCTCTGACCGGACGAGGAGGCGGACCTTG 240
DB 61 IleLeuSerGlnAspArgPheTyrLysValLeuThrAlaGluGlnLysAlaLysAlaLeu 80
QY 241 AAAGAGAGTACATTTTGAACATCCAGATCCCTTTGATATATGATTTGATGACAGACT 300
DB 81 LysGlyGlnTyrAsnPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThr 100
QY 301 CTGAAGAACATGTCGAGAGGCAAAACGCTGAGAGTGCAGACTTATGATTTTGTGACACAC 360
DB 101 LeuLysAsnIleValGluGlyLysThrValGluValProThrTyrAspPheValThrHis 120

```

```

QY 361 TCAGGTTACCAAGACACACGCGTGTACCCCTGCGGAGCGTGGTCTGTTTGAAGGCGATC 420
DB 121 SerArgLeuProIleThrThrValValTyrProAlaAspValValLeuPheGluGlyIle 140
QY 421 TTGGTTTTCACAGCAGGAGATCCGAGACATGTTCCACTCGGCGCTTCTGTCGACAC 480
DB 141 LeuValPheTyrSerIleGlnIleArgAspMetPheHisLeuArgLeuPheValAspThr 160
QY 481 GACTCCGAGCGTCAGGCTGTCTCGAAGATTTCTCCGAGACGTGCGCGGAGGAGACCTG 540
DB 161 AspSerAspValArgLeuSerArgValLeuArgAspValArgValArgGlyArgAspLeu 180
QY 541 GACCAATTCGACGAGATACACCACTTCGTAACCGCGGCTTCGAGAGATTCGCTCG 600
DB 181 GluGlnIleLeuThrIleTyrThrThrPheValLysProAlaPheGluGluPheCysLeu 200
QY 601 CCGACAAAGAGATGATCCGATGATCATCCACGAGAGTGGACATATGTTGCCATC 660
DB 201 ProThrLysLysTyrAlaAspValIleIleProArgGlyValAlaAspMetValAlaIle 220
QY 661 AACCTGATGTCGACGACATCCAGACATTTGATGTCGATTCGAAATGACACCGA 720
DB 221 AsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysLysTrpHisArg 240
QY 721 GAGAGGTCCAAATGGGCGGAGCTACAGCGGACCTTTTCGAGCGGAGGACCACTCGG 780
DB 241 GlyGlySerAsnLysArgSerTyrLysArgThrPheSerGluProLysPheIleProGly 260
QY 781 ATGCTGATCTTGGCAACGCTGACATTTGAGTCCAGAGCAGACCCGAC 831
DB 261 MetLeuThrSerCylLysArgSerHisLeuGluLysSerSerArgProHis 277

RESULT 2
UCK1_MOUSE STANDARD; PRT; 277 AA.
AC P52623;
DT 01-OCT-1996 (Rel. 34, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Uridine-cytidine kinase 1 (EC 2.7.1.48) (UCK 1) (Uridine
DE monophosphokinase 1) (Cytidine monophosphokinase 1).
GN UCK1 OR UMPK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Salivary gland;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heish F.,
RA Diatchenko L., Marsina K., Farmer A.A., Rubin G.M., Hong L.,
RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Rana S.S., McMan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Bosak S.A., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hultky S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fehey J., Helton E., Ketterman M., Madan A., Rodrigues S., Sanchez A.,
RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butcherfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE OF 18-277 FROM N.A.

```



```

Db      184 GlnTyrIleThrPheValIysProAlaPheGlnUGlnUheCysLeuProThrIlysTyr 203
Qy      616 GCCGATGTGTCATCCACAGAGAGTGACAAATATGTTGCCATCAACCTGATGTCAC 675
Db      204 AAlaSpAlIlelleProAGGlyAlaAaSPaSmUeValAlaIleAenUeIleValGln 223
Qy      676 CACATCCAGACATTTGTAATGGTGCATCTGCMAATGACCGAGAGAGGGTCCAAATGG 735
Db      224 HisIleGlnIleAspIleLeuAangly-----Gly 232
Qy      736 CGAGCTACAGCGGACCTTTTGTGAGCCAGGGACCAACCTGGATGTCGACTTGCC 795
Db      233 ProSerIyAsrGlnThr-----AsnGlyCysLeuAanglyTyrThrProSerArg 249
Qy      796 AAACGTCACATTTGAGTGCACAGACAGACCCAC 831
Db      250 LysArgGlnAlaSerGlnSerSerArgProHis 261

RESULT 4
UCK2_MOUSE
ID_UCK2_MOUSE STANDARD; PRT; 261 AA.
AC Q99PM9;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Uridine-cytidine kinase 2 (EC 2.7.1.48) (UCK 2) (Uridine
DE monophosphokinase 2) (Cytidine monophosphokinase 2).
GN UCK2 OR UMPK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10099;
[1]
SEQUENCE FROM N.A.
RX MEDLINE=21203813; PubMed=11306702;
RA Van Rompay A.R., Norda A., Linden K., Johansson M., Karlsson A.;
RT "Phosphorylation of uridine and cytidine nucleoside analogs by two
RT human uridine-cytidine kinases."
RL Mol. Pharmacol. 59:1181-1186(2001).
CC -1- FUNCTION: Phosphorylates uridine and cytidine to uridine
CC monophosphate and cytidine monophosphate. Does not phosphorylate
CC deoxyribonucleosides or purine ribonucleosides. Can use ATP or GTP
CC as a phosphate donor. Can also phosphorylate cytidine and uridine
CC nucleoside analogs such as 6-azauridine, 5-fluorouridine, 4-
CC thiouridine, 5-bromouridine, N(4)-acetylcytidine, N(4)-
CC benzoylcytidine, 5-fluorocytidine, 2-thiocytidine, 5-
CC methylcytidine, and N(4)-anisoylcytidine (By similarity).
CC -1- CATALYTIC ACTIVITY: ATP + uridine = ADP + UMP.
CC -1- CATALYTIC ACTIVITY: ATP + cytidine = ADP + CMP.
CC -1- PATHWAY: Pyrimidine salvage pathway.
CC -1- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF236636; AAK14052.1; -.
DR MGD; MG1:1931744; UCK2.
DR InterPro; IPR006082; PRK.
DR InterPro; IPR006083; PRK URK.
DR InterPro; IPR00764; Uridine_kin.
DR Pfam; PF00485; PRK.1.
DR PRINTS; PR00478; PHRIEUKINASE.
DR PRINTS; PR00988; URIDINKINASE.
DR TIGRFAMs; TIGR00235; udk; 1.
KW transferase; Kinase; ATP-binding.
NP BIND 27 ATP (POTENTIAL).
SEQUENCE 261 AA; 29404 MW; 780AA3C9F5C8B153 CRC64;

```

Alignment Scores:

Prod. No.:	1.09e-62	Length:	261
Score:	932.00	Matches:	186
Percent Similarity:	78.31%	Conservative:	27
Best Local Similarity:	68.38%	Mismatches:	43
Query Match:	61.56%	Indels:	16
DB:	1	Gaps:	4

US-09-896-522-3 (1-834) x UCK2_MOUSE (1-261)

```

QY 22 GACTGCGAGAGCCCGCGCGAGAGCCGACCTCCGAC---CAGCGGCCCTTCTGATA 78
   |||
Db 4 AspSerGIuGIInThrLeuGIInAmHIEGIInGIInProAsnGIyGIyGIInProPheLeuIle 23
QY 79 GGGGTGAGCCCGCCGACATGCCAGGGGAGAAAGTCCACCGTGTGTGAGAAAGATCAAGAGTTG 138
   |||
Db 24 GlyValSerGIyGIyThrAlaSerGIyLysSerSerValLysAlaLysIleValGIInLeu 43
QY 139 CTGGGACAGAGAGGAGTGAAGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 198
   |||
Db 44 LeuGIyGIInAsnGIInValaSerGIyThrLeuGIInValaIleLeuSerGIInAspSer 63
QY 199 TTCTACAGAGTCTCTGAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 258
   |||
Db 64 PheGIyArgValLeuThrSerGIInGIyAlaIyAlaValaLeuGIySGLGIInPheAsnPhe 83
QY 259 GACCATTCAGATGCTTGTATGATTTGATGACAGAGACTGTGAAGAAACATGTGTGAG 318
   |||
Db 84 AspHisIleProAspAlaPheAspAsnGIInIlePheIyThrLeuGIySGLIleThrGIIn 103
QY 319 GGGAAAAGCGTGGAGCGGACCATGATTTGTGACACACTCAAGGATTACAGAGAGAGC 378
   |||
Db 104 GIyGIyThrValGIInIleProValIyAspPheValaSerHisSerArgIySGLGIInThr 123
QY 379 ACGGTGCTACCGTGGAGAGTGTGTGTGAGGAGCATTTGTGTCTACAGAGCAG 438
   |||
Db 124 ValThrIleThrProAlaAspValaIleValaPheGIInIleLeuAlaPheIyTrSerGIIn 143
QY 439 GAAGATCCGGAGACATGTTCCACCTGCGGCTCTGTGTGACACCGACTCCGACGTCAGGCTG 498
   |||
Db 144 GIyValaIyArgAspLeuPheGIInMetIyLeuPheValaIyAspThrAspAlaAspThrArgLeu 163
QY 499 TCTCGAAGAGTCTCCGCGAGAGTGT---CGCGAGAGGAGAGAGAGAGAGAGAGATTCAGC 555
   |||
Db 164 SerArgIyArgValaLeuAspIleSerGIInArgGIyArgAspLeuGIInIleLeuSer 183
QY 556 CAGTACACCACTTCGTGAAGCGGCGCTTCGAGAGAGTTCGCTGCGGACAAAGAACTAT 615
   |||
Db 184 GIInThrIleThrPheValaIySerProAlaPheGIInGIInPheCysLeuProThrIySGLyTyr 203
QY 616 GCGGATGTATCATCCAGAGAGAGTGAACAATATGTTGCCATCAACTGATGTGCGAG 675
   |||
Db 204 AlaAspValaIleIleProArgGIyAlaAspAsnLeuValaIleAsnLeuIleValGIIn 223
QY 676 CACATCCAGAGCATTTGATGGTGAATGTCGCAATGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 725
   |||
Db 224 HisIleGIInAspIleLeuAsnGIyGIyLeuSerIyArgGIInThrAsnGIy----- 240
QY 736 CGGAGCTACAGAGGAGACCTTTTTCGAGCCGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 795
   |||
Db 241 -----TyrLeuAsn-----GlyTyrThrProSerArg 249
QY 796 AAACGGTACATTTGAGAGTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 831
   |||
Db 250 LysArgGIInAlaSerGIInSerSerArgIyProHis 261
  
```

DE Probable uridine-cytidine kinase (EC 2.7.1.48) (UCK) (Uridine
DE monophosphokinase) (Cytidine monophosphokinase).
GN CG6364.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxId=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananthides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers J.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Abmayyan A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballwey R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Burks K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fosler K., Gabrielian A.B., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibeagwam C.,
RA Jalaal M., Kalush F., Kapen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleby J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler P., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun B.,
RA Styriska K., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weisenbach J., Wu Q.A.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhan G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
CC -|- CATALYTIC ACTIVITY: ATP + uridine = ADP + UMP.
CC -|- CATALYTIC ACTIVITY: ATP + cytidine = ADP + CMP.
CC -|- PATHWAY: Pyrimidine salvage pathway.
CC -|- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AE003747; AAF56274.1; -
CC FLYBase: FBgn0039179; CG6364.
CC InterPro: IPR006083; PRK URK.
CC InterPro: IPR00764; uridine_kin.
CC Pfam: PF00485; PRK_1.
CC PRINTS: PR00986; URIDINKINASE.
CC Hypothetical protein; Transferase; Kinase; ATP-binding.
CC NP_BIND 34 41 ATP (POTENTIAL).
SQ SEQUENCE 260 AA; 29213 MW; 2BE57423704925E6 CRC64;

Alignment Scores:

Score: 4.39e+48 Length: 260
 Pred. No.: 738.00 Matches: 143
 Percent Similarity: 82.94% Conservative: 32
 Best Local Similarity: 67.77% Mismatches: 34
 Query Match: 48.75% Indels: 2
 DB: 1 Gaps: 2

US-09-896-522-3 (1-834) x UCK_DROME (1-260)

QY 67 CCCTCTCTGATAGGGGTGAGCGCGGCGCATCTCCAGCGGGAAGTCGACCGTGTGAGAG 126
 Db 27 ProthelenuilegilyvalaiglyglyThrilaSerglysserThralValCyslelyls 46
 QY 127 ATCATGAGTTCGTCGGGACAGAGAGGTGAACAACGCGGACCGGAAGTGTGATCCTG 186
 Db 47 IlenetGuglnleuglyglhlaGlnwetrsphsthrGlnaGlnValaValSerlle 66
 QY 187 AGCCAGGACAGGTTCTACAGAGTCTCGACGACAGGAGGACGAGGACCGCTTGAAAGA 246
 Db 67 SerGlnaspserserPhetyrArGlnleuthrProhlaGlnuylsAlaValaGlnlyGly 86
 QY 247 CAGTACAAATTTGACATCCAGATGCTTGAATAATGATTGATGACAGAGACTCTGAAG 306
 Db 87 LeuhsnphsPheasphlsProasphlaPheasnGlnGlnleuMetTyserThrlengln 106
 QY 307 AACATCGTGAGGCGCAAAACGATGAGGTGAGGAGGCTGATGATTTTGACACACTCA--- 363
 Db 107 AsnilleuylsGlyhslslyValaGlnleProserlyrAspTyArGlnhAsnSerleu 126
 QY 364 AGGTTACGAGACACGCGTGTACCTTCGCGACGCTGCTTCTGTTGAGGCGATCTTG 423
 Db 127 AsphegluValenValleValilleTyrrProhlaAsphValleuPhegluGlylleu 146
 QY 424 GTGTTCTACAGCCGAGAGATCCGGACATGTTCCAGCTCGCTCTTGTGACACCGAC 483
 Db 147 ValPheTyrrPhePolysllearGlnleuPhehlsMetlylsleuPheValaAspThrAsp 166
 QY 484 TCGACGTCAGGCTGTCTCGAAGAGTTCTCGGAGCGTG---CGCCGAGGAGGAGACTG 540
 Db 167 SerasphrArGlnleuAlaArGlnValProArGAsphlleasnGlnuArGlyArGAsphleu 186
 QY 541 GAGCAGATTTCTGACGACGATACACACCTTCGTGAAGCGGCTTCGAGAGTTCTGCTG 600
 Db 187 AspAlaValleuthrGlnTyMetThrPheValysProAlaPhegluGlnuPheCysSer 206
 QY 601 CCGCAAGAAGATATGCCGATGTATCATCCACGAGAGAGTGCACATATGTTGCCATC 660
 Db 207 ProThrylsysPheAlaAsphValilleleProArGlyAlaAspAanThrValaIaile 226
 QY 661 AACCTGATCGTCGACGACATCCAGACACTTCTG 693
 Db 227 AspIleuValleuValhshlsilleGlyGlnleu 237

RESULT 6
 URL1_HUMAN STANDARD: PRT, 548 AA.
 AC O9NWZ5: O9H322:
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DE 28-FEB-2003 (Rel. 41, Last annotation update)
 GN Uridine kinase-like 1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Watanabe K., Kumagai A., Itakura S., Yamazaki M., Tachiro H., Ota T.,
 RA Suzuki Y., Odayashi M., Nishi T., Shibahara T., Tanaka T.,
 RA Nakamura Y., Isoigai T., Sugano S.;
 RT "NEDO human cDNA sequencing project.";

RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21638749; PubMed=11780052;
 RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
 RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,
 RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
 RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
 RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
 RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
 RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,
 RA Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,
 RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
 RA Hammond D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
 RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
 RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
 RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
 RA Lehesvaisto M.H., Leverisa M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
 RA Marsh V.L., Martin S.L., McConachie L.J., McIlroy K., Murray A.A.,
 RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
 RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.T.,
 RA Phillimore B.J.C.T., Prachalingam S.R., Plumb R.W., Ramsay H.,
 RA Rice C.M., Rose M.T., Scott C.E., Sehra H.K., Showman R., Sims S.,
 RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
 RA Spann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
 RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
 RA Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A.,
 RA Wilmink L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
 RA Rogers J.;
 RT "The DNA sequence and comparative analysis of human chromosome 20.";
 RL Nature 414:865-871(2001).
 CC -1- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC -----
 DR EMBL: AK000524; BAA91230.1; -;
 DR EMBL: AL118506; CAC15497.1; -;
 DR HSSP: O26998; 1BD3.
 DR Genew: HGNC:15938; URKL1.
 DR InterPro: IPR006082; PRK.
 DR InterPro: IPR006083; PRK.URK.
 DR InterPro: IPR000764; Uridine_kin.
 DR Pfam: PF00485; PRK.1.
 DR PRINTS: PR00478; PHRBLKINASE.
 DR PRINTS: PR00988; URIDINKINASE.
 DR TIGRfams: TIGR00235; udk.1.
 KW Transferase; Kinase; ATP-binding.
 FT NP BIND 105 112 ATP (POTENTIAL).
 FT CONFLICT 219 228 MISSING (IN REF. 2).
 FT CONFLICT 290 290 N -> D (IN REF. 2).
 SQ SEQUENCE 548 AA; 61139 MW; 0CD03697E02FE7DB CRG64;

Alignment Scores:
 Pred. No.: 1.62e+33 Length: 548
 Score: 545.50 Matches: 111
 Percent Similarity: 64.20% Conservative: 45
 Best Local Similarity: 45.68% Mismatches: 60
 Query Match: 36.03% Indels: 27
 DB: 1 Gaps: 5

US-09-896-522-3 (1-834) x URL1_HUMAN (1-548)

QY 25 TCGGAGAGCCCGCGCGG-----GAGCGGACCGT 54
 Db 65 CysLysSerGlnProLeuLeuArGlnrSerLysArGlnrThrAlaGlyArG 84
 QY 55 CCG-----CACACGCGCCCTTCTGATAGGCGTGAAC 87

```

Db      85  ProProtrpTryrAnsluHISGLYThrInserLySgluaLaphaLalIeGLYLeuGly 104
Qy      88  GGGGGGACTCCGACGGGGAAGTCAACCGTGTGTGAAGAATCATGAGTTCTGGGACAG 147
Db      105  GYGLYSerLaserGlyLysThrThrValAlaArgMetIleIleGluAlaLeu----- 122
Qy      148  AACGAGGTGAACAGCGCGACCGGAGGTGTCTATCTCGAGCCAGACAGGTTTACAG 207
Db      123  ---AspValProTrp-----ValValLeuLeuSerMetAspSerPheTryLys 137
Qy      208  GTCCCTGACGACAGAGCAAGGCGCCTTGAAGAAGACATCAATTTGACCATCCA 267
Db      138  ValLeuThrGlnGlnGlnGlnGlnAlaAlaHisAsnPhenAspPheAspHisPro 157
Qy      268  GATGCTTTGATATGATTTGATGATCAACAGACTTGAAGAATCTGGAGGGCAAAACG 327
Db      158  AspAlaPheAspPheAspLeuIleIleSerThrLeuLysLysGlnGlyLysSer 177
Qy      328  GTGAGGTGGCGGACCTATGATTTGTGACACACTCAAGGTTTACAGACACAGCGTGTG 387
Db      178  ValLysValProLetyrAspPheThrThHisSerArgLysAspTryLysThrLeu 197
Qy      388  TACCTCGGACGCTGTCTGTGTGTGAGGCGCATCTGTGTCTACAGCCAGAGATCCGG 447
Db      198  TyrGlyAlaAsnValIleIlePheGlnGlyLeMetAlaPheAlaAspLysThrLeu 217
Qy      448  GACATGTTTCCACTGCGCCTTCTGTGTGACACAGCATCCGACGTCAGGCTGTCTGAAGA 507
Db      218  GluLeuLeuAspMetLysIlePheValAlaAspHisAspLeuValArgLysValArg 237
Qy      508  GTTCTCGGAGACGCG---CGCGGAGGAGGACCTGAGACGATTCGACGCACTACACC 564
Db      238  LeuArgArgAspIleSerGluArgGlyArgAspIleGlnGlyValIleLysGlnTryAsn 257
Qy      565  ACCTTCTGGAAGCGCGCTTTCGAGAGTTCTGCTGCGGCAAGAAAGTACCGCATGTG 624
Db      258  LysPheValLysProSerPheAspGlnTryIleGlnProThrMetArgLysLeuAlaAsp 277
Qy      625  ATCATCCCAAGAGAGTGAACATATAGTTGCCATCAACTGATCGTGCAGACATCCAG 684
Db      278  ValValProArgLysSerGlyAsnThrValAlaIleAsnLeuIleValGlnHisValHis 297
Qy      685  GACATTCGTG 693
Db      298  SerGlnLeu 300

RESULT 7
URL_MOUSE
ID      URL_MOUSE      STANDARD;      PRT;      548 AA.
AC      Q91YL2;
DT      28-FEB-2003 (Rel. 41, Created)
DT      28-FEB-2003 (Rel. 41, Last sequence update)
DE      15-SEP-2003 (Rel. 42, Last annotation update)
DE      Uridine kinase-like 1.
GN      URKL1.
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Cranialia; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxID=10090;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=22388257; PubMed=12477932;
RA      Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA      Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA      Altschul S.F., Zeeberg B., Buetow K.H., Scheafer C.F., Bhat N.K.,
RA      Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heien F.,
RA      Diatchenko L., Mansina K., Farmer A.A., Rubin G.M., Hong L.,
RA      Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.J., Scheetz T.E.,
RA      Brownstein M.J., Urdin T.B., Toshimaki S., Carninci P., Prange C.,
RA      Bosak S.A., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA      Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huljk S.W.,

```

```

RA      Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA      Pabey J., Halton E., Ketterman M., Madan A., Rodrigues S., Sanchez A.,
RA      Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA      Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA      Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA      Butcherfield Y.S.N., Krzywinski M.I., Skalska U., Smalilus D.E.,
RA      Schnerker A., Schein J.E., Jones S.J.M., Marra M.A.;
RT      "Generation and initial analysis of more than 15,000 full-length
RT      human and mouse cDNA sequences";
RL      Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC      -1- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.
CC      CC
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      CC
DR      EMBL, BC016535; AAH16535.1; -
DR      WGD; MG1:1915806; 111007H10Rik.
DR      InterPro; IPR006082; PRK.
DR      InterPro; IPR006083; PRK URK.
DR      InterPro; IPR00764; Uridine_kin.
DR      Pfam; PR00485; PRK; 1.
DR      PRINTS; PR00476; PHRIKINASE.
DR      PRINTS; PR00988; PHRIKINASE.
DR      TRGFAMs; TRGR00235; udk; 1.
KW      Transferase; Kinase; ATP-binding.
FT      NP BIND 105 112 ATP (POTENTIAL).
SQ      SEQUENCE 548 AA; 60841 MW; 99A5CA23B9BD8525 CRC64;

Alignment Scores:
Pred. No.: 4,59e-33 Length: 548
Score: 539.50 Matches: 110
Percent Similarity: 64.20% Conservative: 46
Best Local Similarity: 45.27% Mismatches: 60
Query Match: 35.63% Indels: 27
DB: 1 Gaps: 5

US-09-896-522-3 (1-834) x URL_MOUSE (1-548)
Qy      25  TGGGAGAGCCCGCGCG-----GAGCGGACCGT 54
Db      65  CysLysSerLeuProLeuLeuArgThrSerLysArgThrIleTryrAlaGlyArg 84
Qy      55  CCG-----CACACGCGCCTTCTCGATAGGGGTGAGC 87
Db      85  ProProtrpTryrAnsluHISGLYThrInserLySgluaLaphaLalIeGLYLeuGly 104
Qy      88  GGGGGGACTCCGACGGGGAAGTCAACCGTGTGTGAAGAATCATGAGTTCTGGGACAG 147
Db      105  GYGLYSerLaserGlyLysThrThrValAlaArgMetIleIleGluAlaLeu----- 122
Qy      148  AACGAGGTGAACAGCGCGACCGGAGGTGTCTATCTCGAGCCAGACAGGTTTACAG 207
Db      123  ---AspValProTrp-----ValValLeuLeuSerMetAspSerPheTryLys 137
Qy      208  GTCCCTGACGACAGAGCAAGGCGCCTTGAAGAAGACATCAATTTGACCATCCA 267
Db      138  ValLeuThrGlnGlnGlnGlnGlnAlaAlaHisAsnPhenAspPheAspHisPro 157
Qy      268  GATGCTTTGATATGATTTGATGATCAACAGACTTGAAGAATCTGGAGGGCAAAACG 327
Db      158  AspAlaPheAspPheAspLeuIleIleSerThrLeuLysLysGlnGlyLysSer 177
Qy      328  GTGAGGTGGCGGACCTATGATTTGTGACACACTCAAGGTTTACAGACACAGCGTGTG 387
Db      178  ValGlnValProLetyrAspPheThrThHisSerArgLysLysAspTryLysThrLeu 197
Qy      388  TACCTCGGACGCTGTGTGTGTGAGGCGCATCTGTGTCTACAGCCAGAGATCCGG 447

```

Db 198 TyrGlyAlaAsnValIleIlePheGluGlyIleMetAlaPheAlaAspIleThrLeu 217

Qy 448 GACATGTTCCACCTGCGCTCTGTGGACACCGACCTCGACGTCGTCTCGAAGA 507

Db 218 GluLeuLeuAspMetLysIlePheValAspThrAspSerAspIleArgLeuValArgArg 237

Qy 508 GTTCTCCGGAAGCTG---CGCGAGAGGAGGACCTGGACAGATTCTGACCCAGTACAC 564

Db 238 LeuArgAspAspIleSerGluArgGlyArgAspIleGluGlyValIleGlyGlnThrAsn 257

Qy 565 ACCTTCGTGAAGCGGCTTGGAGAGATTCTGCTCCGACGACAAAGACTATGCCGATGTG 624

Db 258 LysPheValLysProAlaPheAspGlnTyrIleGlnProThrMetArgLeuAlaAspIle 277

Qy 625 ATCATCCAGACGAGGTGACCAATATGTTGCCATCATCGATCGTGCAGACATCCAG 684

Db 278 ValValProArgGlySerGlyAsnThrValAlaIleAspLeuIleValGlnHisValHis 297

Qy 685 GACATTCTG 693

Db 298 SerGlnLeu 300

RESULT 8

UCL_CABEL STANDARD; PRT; 229 AA.

ID UCL_CABEL

AC Q17413:

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DE 28-FEB-2003 (Rel. 41, Last sequence update)

DE Probable uridine-cytidine kinase (EC 2.7.1.48) (UCL) (Uridine

DE monophosphokinase) (Cytidine monophosphokinase).

GN B0001.4.

OS Caenorhabditis elegans.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;

OC Rhabditidae; Peloderinae; Caenorhabditis.

OX NCBI_TaxID=6239;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Bristol N2;

RA Sims M.;

RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.

CC -1- CATALYTIC ACTIVITY: ATP + uridine = ADP + UMP.

CC -1- CATALYTIC ACTIVITY: ATP + cytidine = ADP + CMP.

CC -1- PATHWAY: Pyrimidine salvage pathway.

CC -1- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

CC

DR EMBL; Z69634; CA93453.1; -.

DR PIR; T18629; T18629.

DR WormPep; B0001.4; CE05142.

DR InterPro; IPR006082; PRK.

DR InterPro; IPR006083; PRK URK.

DR InterPro; IPR00764; Uridine_kin.

DR Pfam; PF00485; PRK; 1.

DR PRINTS; PRO0478; PHRIEKLINASE.

DR PRINTS; PRO0988; URIDINKINASE.

KT Hypothetical protein: Transferase; Kinase; ATP-binding.

FT NP_BIND 15 22 ATP (POTENTIAL).

SO SEQUENCE 229 AA; 26195 MW; 890CB21596A820FCC CRC64;

Alignment Scores:

Pred. No.: 4.03e-31 Length: 229

Score: 512.50 Matches: 105

Percent Similarity: 69.05% Conservative: 40

Best Local Similarity: 50.00% Mismatches: 58

Query Match: 33.85% Indels: 7

DB: 1 Caps: 4

US-09-896-522-3 (1-834) x UCL_CABEL (1-229)

Qy 67 CCCTTCTGATAGGGGTGAGCGCGCGACATCCAGCGGGAAGTGCACCGTGTGAGAAG 126

Db 8 ProLeuLeuIleGlyValAlaGlyGlyThrSerCysGlySerThrIleValGluArg 27

Qy 127 ATCATGAGAGTGTGGGACAGACAGAGTGAACAG---CGGACGCGAAGGTGTCATC 183

Db 28 IleIleGluAsnLeuAsnAlaAsnAlaLysGlnSerCylArgGlnIleAspIleValHis 47

Qy 184 CTGAGCCAGACAGAGTGTCTCAAGATCTTACGCGACGACGACGACGACGACGACGACG 243

Db 48 LeuSerLeuHisSerPheTyrArgGluLeuSerAlaGluGlyIleLeuAlaArgGlu 67

Qy 244 GACAGATCAATTTTGCACCATCCAGATGCTTTGATGATGATTTGATGATGATGATG 303

Db 68 GlyLysPheAsnPheAspHisProAspGlnIleAsnPheAspLeuAlaGluThrLeu 87

Qy 304 AAGAATCTCTGAGAGGCGAAAGCGTGGAGTCCGACCTATGATTTTGACACACTCA 363

Db 88 GlnAsnMetIleAspGlyLysThrValGluIleProLysTyrAspMetIleThr---Ser 106

Qy 364 AGGTACCAAGACACACGCGTGTCTACCCCTGCGAGCGTGTCTGTTGAGCGCATCTTG 423

Db 107 SerMetAsnGlyThrValThrValGluProAlaLysValIleIleIleGluGlyIleLeu 126

Qy 424 GTGTTCTACAGCAGAGATCCGAGACATGTCACCTGCTTTCTGTGACACCGAC 483

Db 127 LeuLeuTyrAspGluArgValArgLysLeuLeuSerThrLysLeuPheValGluTyrAsn 146

Qy 484 TCCGAGCTGACGCGTCTCGAAGATT-----CTCGGAGACGTGCGCGAGGAGG 534

Db 147 AlaGluSerArgLeuArgAsnArgLeuAlaThrTyrIleArgAspTyrHisAlaGluPro 166

Qy 535 GACCTGAGCAGATCTGACGACGATACACCACTTCGTGTAAGCGGCTTGCAGAGATTG 594

Db 167 -----LeuSerIleIleArgGlnTyrThrGluPheValLysProAlaPheGluGluPhe 184

Qy 595 TGCCTGCCGCAAAAGAAATGTGCCGATGTGATATCCACGAGAGTGCACATATGTTT 654

Db 185 CysArgProThrLysTyrAlaAspValIleIleProArgLysAlaAspAsnHisVal 204

Qy 655 GCCATCAACCTGATCGTGCAGACATCCAG 684

Db 205 AlaThrAspLeuIleAlaLysAsnLeuGln 214

RESULT 9

URK_BACHD STANDARD; PRT; 211 AA.

ID URK_BACHD

AC Q9KDB8:

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DE Uridine kinase (EC 2.7.1.48) (Uridine monophosphokinase) (Cytidine

DE monophosphokinase).

GN URK OR BH1275.

OS Bacillus halodurans.

OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.

OX NCBI_TaxID=86665;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C-125 / JCM 9153;

RX MEDLINE=20512582; PubMed=11058132;

RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,

RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,

RA Horikoshi K.;

RT "Complete genome sequence of the alkaliphilic bacterium Bacillus

RT halodurans and genomic sequence comparison with Bacillus subtilis.";

RL Nucleic Acids Res. 28:4317-4331(2000).

CC -1- CATALYTIC ACTIVITY: ATP + uridine = ADP + UMP.

CC -1- CATALYTIC ACTIVITY: ATP + cytidine = ADP + CMP.

CC -1- PATHWAY: Pyrimidine salvage pathway.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

 DR EMBL/ AP001511; BAB04394.1; -
 DR PIR/ C83809; C83809.
 DR HAMAP/ MF_00551; -1.
 DR InterPro/ IPR003593; AAA_ATPase.
 DR InterPro/ IPR006082; PRK.
 DR InterPro/ IPR006083; PRK URK.
 DR InterPro/ IPR00764; Uridine_kin.
 DR Pfam/ PF00485; PRK; 1.
 DR PRINTS/ PRO0478; PHRIBLKINASE.
 DR PRINTS/ PRO0988; URIDINKINASE.
 DR TIGRPFAM/ TIGR00235; udk; 1.
 DR Transferrase/ Kinase; ATP-binding; Complete proteome.
 DR KW/ NE BIND 12 19 ATP (POTENTIAL).
 DR FT/ NE BIND 12 19 ATP (POTENTIAL).
 DR SQ SEQUENCE 211 AA; 24387 MW; C2AF82CB030520B CRC64;

Alignment Scores:
 Pred. No.: 3,73e-28 Length: 211
 Score: 473.00 Matches: 93
 Percent Similarity: 66.51% Conservative: 50
 Best Local Similarity: 43.26% Mismatches: 56
 Query Match: 31.24% Indels: 16
 DB: 1 Gaps: 4

US-09-896-522-3 (1-834) x URK_BACHD (1-211)

QY 61 CAGCGGCGCTTCCTGATGAGGAGTGGAGCTGCGAGCGGGAAGTGCAGCTGTGT 120
 DB 3 LysArgProIleIleIleGlyValAlaGlyGlyThrIleSerGlyLysThrIleValAla 22
 QY 121 GAGAAGATCATGAGTGTCTGCGACAGAAAGAGGTGGAAGCGCGCGAGAGTGTGTC 180
 DB 23 LysGluIlePhe-----TyrGlnPheGlnGluLysSerIleVal 35
 QY 181 ATCTTGAGCCAGACAGGTCTTCAAG-----GTCTTGACGCGAGACAGAAAGGCC 231
 DB 36 LcuIleGluIleAspAlaTyrTyrLysAspGlnSerGlnLysSerIleGluIleArgLcu 55
 QY 232 AAGGCGCTGAAAGAGCATATTTTGAACATCCAGATCCCTTGTATATGATTTGATG 291
 DB 56 -----GlnThrAsnTyrAspHisProLeuAlaPheAspAsnAspLeuLcu 70
 QY 292 CACAGGACTGTGAAGAATCGTGTGAGGCGGAAACGGTGTGAGGTGCGACCTATGATTTT 351
 DB 71 IleGluIleIleuHisSerIleuLeuAenglyGlnAlaIleIleuLysProValTyrAspTyr 90
 QY 352 GTACACACTCAAGGTTTACAGAGACCAAGTGTGTTACCTTGGCGAGCTGTGTTGTTT 411
 DB 91 LysLeuHisThrArgSerAsnGluValIleLeuValGluProLysAspValIleIleLcu 110
 QY 412 GAGGCGCATCTTGTTGTTCTACACCGAGAGATCCGGAGACATGTTCCACCTGCCCTTTC 471
 DB 111 GlnGlyIleLeuLeuLeuGluAspGluArgLeuArgGluLysSerIleLysLeuPhe 130
 QY 472 GTGAGACACGACTCCGAGCTGTGCTGTGAGAGATTTCCGGGAGCTGTGCC--CGA 528
 DB 131 ValAspThrAspAlaAspIleArgIleIleIleArgHisMetAlaArgAspIleArgGluArg 150
 QY 529 GGAAGGAGCTGTGAGAGATTTCTGACGACATCAACCACTTGTGAAGCCGGGCTTCGAG 588
 DB 151 GlyArgThrIleuGlnSerValIleGluGlnTyrThrLysValValArgProMetHisMet 170

QY 589 GAGTCTGCTGCGCGACAAAGATGCGGATGATCATCCACGAGAGTGCAGCAAT 648
 DB 171 GlnPheIleGluProThrIleArgTyrAlaAspValIleIleProGluIleGlnAsn 190
 QY 649 ATGGTTGCATCAACTGATTCGTGACGACATCCAGACATTCGTG 693
 DB 191 ArgValAlaIleAspLeuMetValThrLysIleArgAlaIleIle 205

RESULT 10
 URK_CLOPE
 ID URK_CLOPE STANDARD; PRT; 208 AA.

AC Q8XJ16;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Uridine Kinase (EC 2.7.1.48) (Uridine monophosphokinase) (Cytidine
 DE monophosphokinase).
 GN UDK OR CPE1770.
 OS Clostridium perfringens.
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
 OC Clostridium.
 ON NCBI_TaxId=1502;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=J3 / Type A;
 RX MEDLINE=21664373; PubMed=11792842;
 RA Shimizu T., Ohtani K., Hirakawa H., Ohshima K., Yamauchi A.,
 RA Shiba T., Ogasawara N., Hattori M., Kohara S., Hayashi H.;
 RT "Complete genome sequence of Clostridium perfringens, an anaerobic
 RT flesh-eater.";
 RT Proc. Natl. Acad. Sci. U.S.A. 99:996-1001(2002).

CC -1- CATALYTIC ACTIVITY: ATP + uridine = ADP + UMP.
 CC -1- CATALYTIC ACTIVITY: ATP + cytidine = ADP + CMP.
 CC -1- PATHWAY: Pyrimidine salvage pathway.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

 DR EMBL/ AP003191; BAB81476.1; -
 DR HAMAP/ MF_00551; -1.
 DR InterPro/ IPR006082; PRK.
 DR InterPro/ IPR006083; PRK URK.
 DR InterPro/ IPR00764; Uridine_kin.
 DR Pfam/ PF00485; PRK; 1.
 DR PRINTS/ PRO0478; PHRIBLKINASE.
 DR PRINTS/ PRO0988; URIDINKINASE.
 DR TIGRPFAM/ TIGR00235; udk; 1.
 DR Transferrase/ Kinase; ATP-binding; Complete proteome.
 DR KW/ NE BIND 11 18 ATP (POTENTIAL).
 DR FT/ NE BIND 11 18 ATP (POTENTIAL).
 DR SQ SEQUENCE 208 AA; 23884 MW; 941BC368740DB5E CRC64;

Alignment Scores:

Pred. No.: 8.84e-28 Length: 208
 Score: 468.00 Matches: 88
 Percent Similarity: 66.51% Conservative: 55
 Best Local Similarity: 40.93% Mismatches: 56
 Query Match: 30.91% Indels: 16
 DB: 1 Gaps: 4

US-09-896-522-3 (1-834) x URK_CLOPE (1-208)

QY 61 CAGCGGCGCTTCCTGATGAGGAGTGGAGCTGCGAGCGGGAAGTGCAGCTGTGT 120
 DB 2 LysArgProIlePheIleIleGlyValAlaGlyGlyThrIleSerGlyLysThrIleValAla 21
 QY 121 GAGAAGATCATGAGTGTCTGCGACAGAAAGAGGTGGAAGCGCGAGAGTGTGTC 180

Db 22 LysGluIleTyrArgGlnPheGlyGluAsp-----CysIleAla 34
161 ATCTCGAGCCAGGAGCTTCTTACAAG-----GTCTCGACGACAGACGAAGGCC 231
35 MetIleGluGlnAspSerTyrTyrLysAsnGlnSerHisIleuSerMetGluAspArgVal 54
232 AAGGCGCTTGAAGAAGACAGTACAATTTTGAACCATCCAGATGCCCTTTATTAATGATTGGAT 291
55 LysThr-----AsnTyrAspHisProAsnAlaPheAspAsnIleuLeu 69
292 CACAGGACTCTGAGAACAATCGGAGGGGCAAAACGGTGAAGGTGCCGACTATGATTTT 351
70 ValSerHisLeuGlnSerIleuAsnGlnHisSerIleGlnLysProSerTyrAspPhe 89
352 GTGACACACTGACAGGTTTACAGAACCAACGATGCTTACCTCGGACGCTGGTTCGTT 411
90 SerIleHisAsnArgIleGluAspThrThrLysValGlnProLysGluIleValIleVal 109
412 GAGGCGATTGTGTGTTCTTACACGCCAGAGATTCGGGACATGTTCCACTCGGCTCTTC 471
110 GlnGlyIleLeuIleIleuGlnAspProArgIleArgGluIleuAspIleLysIleTyr 129
472 GTGACACCGGACTCGGACGCTGCGTCTCGAAGAGTCTCCGGGACGTC---CGCCGA 528
130 ValAspThrAspAlaAspValArgIleIleArgGlnMetValArgAspIleAsnGluArg 149
529 GGGAGGACCTTGAGACAGATTCGTGACGCGACTACACCACTTGCTGAAGCGGCTTCGAG 588
150 GlnArgThrMetGlnSerValIleAsnGlnTyrLeuAsnValValLysProMetHisAsn 169
589 GAGTTCTGCTCCCGACAAAGATGTATGCCGATGTGATCATCTCCACGAGAGTGACAAT 648
170 GlnPheThrGlnProThrLysLysPheAlaAspIleIleIleProGluGlnGlyHisAsn 189
649 ATGGTTCGATCAACCTGATCGTCGACGACATCCAGGACATTTCTG 693
190 LysValAlaIleAspIleValAlaLysIleLysGluValIleu 204

RESULT 11
URK_LACLA
ID URK_LACLA STANDARD; PRT; 206 AA.
AC 09CFE21;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Uridine kinase (EC 2.7.1.48) (Uridine monophosphokinase) (Cytidine
monophosphokinase).
GN UDK OR LIL1660.
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
OX NCBI_TaxID=1360;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IL1403;
RX MEDLINE=21235186; PubMed=11337471;
RA Bojotin A., Wincker P., Manger S., Jallion O., Malarme K.,
Weissenbach J., Ehrlich S.D., Sorokin A.;
RT The complete genome sequence of the lactic acid bacterium Lactococcus
lactis ssp. lactis IL1403.";
RL Genome Res. 11:731-753 (2001).
CC -1- CATALYTIC ACTIVITY: ATP + uridine = ADP + UMP.
CC -1- CATALYTIC ACTIVITY: ATP + cytidine = ADP + CMP.
CC -1- PATHWAY: Pyrimidine salvage pathway.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>)

CC	or send an email to license@isb-sib.ch).	
CC	-----	
DR	EMBL; AE006396; AAK05758.1; -	
DR	PIR; D86832; D86832.	
DR	HAMAP; MF_00551; -; 1.	
DR	InterPro; IPR006082; PRK.	
DR	InterPro; IPR006083; PRK URK.	
DR	InterPro; IPR00764; Uridine_kin.	
DR	Pfam; PF00485; PRK; 1.	
DR	PRINTS; PR00478; PHRIBLKINASE.	
DR	PRINTS; PR00988; URIDINKINASE.	
DR	TIGRFAMs; TIGR00235; udk; 1.	
KW	Transferase; kinase; ATP-binding; Complete proteome.	
FT	NP_BIND 11	18 ATP (POTENTIAL).
SQ	SEQUENCE 206 AA; 23740 MW; 8238BF052037DBEE CRC64;	
Alignment Scores:		
Pred. No.:	9,996-27	Length: 206
Score:	454.00	Matches: 92
Percent Similarity:	62.50%	Conservative: 43
Best Local Similarity:	42.59%	Mismatch: 65
Query Match:	29.99%	Indels: 16
DB:	1	Gaps: 5
US-09-896-5522-3 (1-834) x URK_LACLA (1-206)		
QY	61 CAGCGGCCCTTCCTATAGAGGTGAGCGCGCGACTGCCAGCGGAAATCGACCGTGTG	120
DB	2 Lyslysthrleulleilleglyalthrghlyglyserlasercllysthrsevalser	21
QY	121 GAGAAATCATGAGTCTCTGGGACAGAACGAGGTGAAACAGCGGACCGAAGTGTG	180
DB	22 HIsalalleuGlnthrPhe---SerAsnGln-----Arglleala	34
QY	181 ATCTGACCCAGAGAGGTCTACAG-----GTCTGACCGACAGACGAAGGCC	231
DB	35 MetlleGluHIsAspSerlyrlyrlylaspGlnserHIsleuthrPheGlnGluAlaGthr	54
QY	232 AAGCGCTTGAAGGACAGTACATTTTGACCCATCGAGATGCTTTGATGATGATTGATG	291
DB	55 LysThr-----AsnlyrAspHIsProlleuAlaPheAspThrAspThrleu	69
QY	292 CACAGACTCTGAAGACATGCTGAGGAGCAAAACGCTGAGAGTGCACCTATGATTT	351
DB	70 llealaglnleuLysGlnleuGlnlyrGlyrAlaValaIAspIleProlleThrAspThr	89
QY	352 GTGACACACTGAAGTTACCAAGAGCCACACGCGTGTCTACCCGTGGACGCTGTTCTGTT	411
DB	90 AlaIysHIsThrAspSerGlnGlnlyrlyrArgGlnGlnProvalaIAspAlaIleval	109
QY	412 GAGGCACTTGTGTGTTCTTACAGCCAGAGAGATCCGAGCATGTTCCACTGCGCTCTTC	471
DB	110 GlnGlylleleuValleuGlnAspGlnlyrGleuAlaGlyAspIleuMetAspIlelyllePhe	129
QY	472 GTGACACCGACTCCGACGTCAGAGCTGTCTCGAAGAGTTCTCCGGGACGTG---CGCGA	528
DB	130 ValAspThrAspAspValArglleThrArgGlylleArgArgAspIleGlnGlnlyr	149
QY	529 GCGAGGACCTGAGACAGATTTCTGACGACAGTACACCACTTCTGTGAAGCGCGCTTCGAG	588
DB	150 GlyrlyrThrleuAspSerValIleThrGlnlyrleuAspAlaValIlyPhePrometlyrHIs	169
QY	589 GAGTTTCGCTCCGCGACAAAGAGATGCCGATGATCATCCACGAGAGTGGACAT	648
DB	170 GlnPheIleGlnProthrLysArglyrAlaAspValIlellePheGlnGlnlyrAlaSerAsn	189
QY	649 ATGTTGCATCAACCTGATGCTGACGACACATCCAGACATCTGAT	696
DB	190 ThrValGlyValAspIlelleThrThrlylleAlaSerlleleuAsn	205

AC Q99TN8;
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Uridine kinase (EC 2.7.1.48) (Uridine monophosphokinase) (Cytidine monophosphokinase).
 GN UKR OR SAV1611 OR SA1439 OR MW1561.
 OS Staphylococcus aureus (strain M50 / ATCC 700699),
 OS Staphylococcus aureus (strain N315), and
 OS Staphylococcus aureus (strain MW2).
 OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
 OX NCBI_TaxID=158878, 158879, 196620;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=M50 / ATCC 700699, and N315;
 RX MEDLINE=21311952; PubMed=11418146;
 RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
 RA Kato L., Oguchi A., Aoki K.-I., Nagai Y., Iian J.-O., Ito T.,
 RA Kanamori M., Matsunaru H., Maruyama A., Murakami H., Hoshoyama A.,
 RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
 RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
 RA Kanihisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
 RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
 RT "Whole genome sequencing of methicillin-resistant Staphylococcus aureus."
 RL Lancet 357:1225-1240(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MW2;
 RX MEDLINE=22040717; PubMed=12044378;
 RA Nagai T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
 RA Nagai Y., Iwama N., Amano K., Naimi T., Kuroda H., Cui L.,
 RA Yamanoto K., Hiramatsu K.;
 RT "Genome and virulence determinants of high virulence community-acquired MRSA."
 RL Lancet 359:1819-1827(2002).
 CC -|- CATALYTIC ACTIVITY: ATP + uridine = ADP + UMP.
 CC -|- CATALYTIC ACTIVITY: ATP + cytidine = ADP + CMP.
 CC -|- PATHWAY: Pyrimidine salvage pathway.
 CC -|- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -|- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.ebi.ac.uk/announcements> or send an email to license@ebi.ac.uk).
 CC
 CC EMBL, AP003362; BAB57773.1; -
 DR EMBL, AP003134; BAB42703.1; -
 DR EMBL, AP004827; BAB95426.1; -
 DR PIR, B89943; B89943.
 DR HAMAP, MF_00551; -; 1.
 DR InterPro: IPR006082; PRK.
 DR InterPro: IPR006083; PRK.
 DR InterPro: IPR00764; Uridine_kin.
 DR Pfam, PF00485; PRK; 1.
 DR PRINTS, PR00478; PHRIKINASE.
 DR PRINTS, PR00988; URIDINKINASE.
 DR TIGRFam, TIGR00235; udk; 1.
 KW Transferase; Kinase; ATP-binding; Complete proteome.
 FT N_BIND 11
 FT S_Q 207 AA; 23505 MW; FD3A6D4E531DBD CRC64;
 Alignment Scores:
 Pred. No.: 3 36e-26 Length: 207
 Score: 447.00 Matches: 88
 Percent Similarity: 65.09% Conservative: 50
 Best Local Similarity: 41.51% Mismatches: 58
 Query Match: 29.52% Indels: 16
 DB: 1 Gaps: 4

US-09-896-522-3 (1-834) x URK_STAM (1-207)
 QY 73 CTGATATGGGTGAGCGCGGCGACTGCCAGCGGAGATCGACCTGTGTGAGAAATCATG 132
 Db 6 IIEIIEGLIYIEAIGAGIYISERGILYSERGILYSETHRVALTHIRANGIUIIEMET 25
 QY 133 GAATTCTGGGACAGAACAGATGGAACAGCGGACCGGAGAGGTGTGATCTCGAGCCAG 192
 Db 26 LYSANLEU-----GlutylHisSerValAlaLeuLeuAlaGln 38
 QY 193 GAGAGCTTCTAC-----AAGTCTGACGCGAGAGAGGACCAAGCCTTGA 243
 Db 39 AspTyrTrpTyrTrpAspGlnTyrShiSleThrPhaSpGluArgLeuGln----- 56
 QY 244 GAGACGATACATTTGATCCATCCATGATGCTTGTATGATGATTTGATGACAGACTTG 303
 Db 57 -----AsnTyrAspHisProPheAlaPheAspAsnAspLeuLeuIleGluAsnLeu 73
 QY 304 AAGAACATCGTGAAGGCAAAACGTCGAGAGTCGCGACCTATGATTTGTGACACACTCA 363
 Db 74 LYSAspLeuLysAsnGlyLysAlaValGluValProThrTyrAspTyrAlaSerHisThr 93
 QY 364 AGCTTACGAGACGACGCGTGTCTACCTGCGACGTCGATGCTTGTGAGGATCTTG 423
 Db 94 ArgSerAspIleThrIleAspPheLysProLysAspValIleIleValGluGlyIlePhe 113
 QY 424 GTTCTTACACCGAGAGATCCGGAATGTTCCACTGCGCTTCTGACACCGAC 483
 Db 114 AlaLeuGluAsnLysValIleValArgAspMetMetAspValLysIleTyrValAlaSerThrAsp 133
 QY 484 TCCGAGTCAGCGTCGTCTCGAAGATTCCTCGGAGAGCTGGC---GAGGAGAGAGACTG 540
 Db 134 AlaAspLeuArgIleLeuLysArgLeuThrArgAspThrLysGluArgGlyAspSerMet 153
 QY 541 GAGCAATTCGTGACGAGTACACACCTTCGTGAAGCCGCGCTTCGAGAGTTCGCTG 600
 Db 154 AspSerValIleAsnGlnTyrLeuSerValValArgProMetHisAspGlnPheIleGlu 173
 QY 601 CCGACAAAGAGTATGCCGATGATCATTCCTCCAGAGAGTGGACAAATATGTTGCCATC 660
 Db 174 ProThrLysLysThrLysIleAspIleIleIleProGluGlyGlySerAsnLysValAlaIle 193
 QY 661 AACCTATGTCGACGACATCCAGCATTCGTAAT 696
 Db 194 AspIleMetThrThrLysIleGlnSerLeuValSer 205
 RESULT 13
 URK_STRP8
 ID URK_STRP8 STANDARD; PRT; 208 AA.
 AC Q8P0F8;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Uridine kinase (EC 2.7.1.48) (Uridine monophosphokinase) (Cytidine monophosphokinase).
 GN UKR OR SPYMI8_1380.
 OS Streptococcus pyogenes (serotype M18).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=186103;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MGAS8232 / Serotype M18;
 RX MEDLINE=21927593; PubMed=11917108;
 RA Smoot J.C., Barbican K.D., Van Gompel J.J., Smoot L.M., Chaussee M.S.,
 RA Sylva G.L., Sturdevant D.E., Ricklefs S.M., Porcella S.F.,
 RA Parkins L.D., Berez S.B., Campbell D.S., Smith T.M., Zhang Q.,
 RA Kapur V., Daly J.A., Veasey L.G., Nusser J.M.;
 RT "Genome sequence and comparative microarray analysis of serotype M18 Group A Streptococcus strains associated with acute rheumatic fever outbreaks."
 RT Proc. Natl. Acad. Sci. U.S.A. 99:4668-4673(2002).

Qy	18	ATCTGAGCCAGGACGAGTTTCAAGAGGCTCGACGGACGACGAGGCGCAAGGCGCTTG	240
Db	95	ILEuSerMetcAspSerPheTyrIysValIleuThrProGluGluIleIysAlaIleHis	114
Qy	241	AAAGCAGTACCAATTTTGCATCCATCCAGATGCGCTTGGATTAATGATTGTGATCAGAGACT	300
		::: ::	
Db	115	GIuSerArgTyrAsnPheAspGlyProAsnAlaPheAspPheAspLeuTyrGluVal	134
Qy	301	CTGAGAACATCGTGGAGGGCAAAACGGTGGAGGCGCGACCTTATGATTTGTGACAC	360
Db	135	LeuIysArgLeuArgGluGlyIysSerValAspValProValTyrAspPheAsnThrHis	154
		::: ::	
Qy	361	TCAAGTTTACCAAGACCAACGGTGATCTACCCCTGGCGACGTGTTCTGTTGAGGGATC	420
Db	155	SeArgAspProAsnSerIysMetMetTyrGlyAlaAspValLeuIlePheGluGlyIle	174
		::: ::	
Qy	421	TTGGTGTTCACAGCCAGGAGATCCGGGACAGTTCCACCTGGCGCTTCTGTGACAC	480
Db	175	LeuAlaPheHisAspGluIleuArgIleIysValAsnLeuMetAspMetLysValAlaPheValAspThr	194
Qy	481	GACTCCGACGCTCAGGCTGTCTCGAAGAGTTCTCGGACGTCGCGC--CGAGGAGGAGAC	537
Db	195	AspGlyAspPheuArgLeuAlaArgArgIleValArgAspValThrAspArgGlyArgAsp	214
Qy	538	CTGAGCAGATTTCTGACGCGATGACACCACTTCGTGAAGCGCGGCTTCGAGGATTCGC	597
Db	215	ILEAspGlyIleMetGluGlnTyrPheThrPheValLysProAlaPheAspLysTyrIle	234
Qy	598	CTGCGCAGAAAGATGATCCGATGTGATCATCCACAGAGAGTGGACAATATGTTGCC	657
Db	235	AlaProCysMetcAspSerAlaAspLeuIleValProArgGlyGlyGluAsnAspValAla	254
Qy	658	ATCAACTGATCTGCGACGACATC	681
Db	255	ILEAspMetIleValGlnAsnVal	262

```

RESULT 2
T21110
hypothetical protein F19B6.1b - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C/Accession: T21110
R/Thomas, K.
submitted to the EMBL Data Library, February 1996
A/Reference number: Z19375
A/Accession: T21110
A/Status: Preliminary; translated from GB/EMBL/DBDJ
A/Molecule type: DNA
A/Residues: 1-555 <M11>
A/Cross-references: EMBL:Z69635, NID:g1200023; PIDN:CAA93462.1; GSPDB:GN00022; CESP:F19B6
A/Experimental source: clone F19B6
A/Gene: CESP:F19B6.1b
A:Map position: 4
A:introns: 35/1; 83/3; 201/3; 267/3; 395/2; 511/3

Alignment Scores:
Pred. No.:          Length:          555
Score:              559.00           110
Percent Similarity: 74.04%           Matches:
Best Local Similarity: 52.88%         Conservative: 44
Query Match:        36.92%           Mismatches: 46
DB:                  Indels:          8
                        Gaps:          2

US-09-896-522-3 (1-834) x T21110 (1-555)

OY      61  CAGCGGCCCTTCTGATGAGGGGTGAGCGCGGCACTGCCAGCGGGAAGTCCAGCGTGTGT 120
      :: |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      102  LysHISpPropheValIIeGIyValCysGIySerAlaSerGIyLysThrValAla 121

OY      121  GAGAAAGATCAATGAGTACTGTGGGACAGAAACGAGGTGGAACACGCGGACGGAAGTGTCT 180
      |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      122  GluIySrlIleValIGluArgLeuGly-----IleProTrpValThr 134

```

QY ATCCGAGCAGGACGAGGTTCCTACAAGGCCCTGACGGCAGACAGCAAGGCCTTGA 240
181 |||||
Db IleuSerMetAapSerPheTyrIleValLeuThrProGIuGlutIleLysAlaIaHis 154
135 |||||
QY AAAGACAGTACAATTTTGACCATTCCAGATGCGCTTTGATATGATTGTGACACAGACT 300
241 :::::
Db GluSerArgTyrAsnPhaAspPolYProAsnAlaPheAspPheAspLeuTyrGIuVal 174
155 |||||
QY CTGAAGAACAATCGTAGAGGGCAAACGGGAGGAGCGCACCTTAATGATTTTGACACAC 360
301 |||||
Db LeuLysAagLeuAlaGluGIlyLysSerValAspValProValTyrAspPheAsnThrHis 194
175 |||||
QY TCAAGATTACAGAGACAGCGGTGCTAACCCCTGCGGACGTGATTCCTGTGAGGCGCATC 420
361 |||||
Db SerArgAspProAsnSerLysMetMetCtyrGIyAlaAspValLeuIlePheGluGIyIle 214
195 |||||
QY TTGGTGTTCTACAGCCAGAGAGATCTCCGGACATGTTCCACTCGCGGCTCTTCGGGACACC 480
421 |||||
Db LeuAlaPheHisAaspGIuArgIleLysAsnLeuMetAspMetLysValPheValAspThr 234
215 |||||
QY GACTTCGACGTCAGGCGTGTCTTCGAAAGATTCCTCGGACGTGCGC---CGAGGAGGAGAC 537
481 |||||
Db AspGIyAspLeuArgLeuAlaArgIleValAlaArgAspValThrAspArgGIyArgAsp 254
235 |||||
QY CTGAGACGAGATTCTACGCACTTACACCACTTCCTCGGAACCGCGCCTTCGAGAGATTCTGC 597
538 |||||
Db IleaspGIyIleMetGluGlnTyrPheThrPheValLysProAlaPheAspLysTyrIle 274
255 |||||
QY CTGCCGACAAAGAAATATGATCCGATGTGATCATCCCACGAGAGTGAACATAATGTGTGCC 657
598 |||||
Db AlaProCysMetAaspSerAlaAspLeuIleValProArgGIyGIyGluLysAspValAla 294
275 |||||
QY ATCAACCTGATGTGACGACATC 681
658 |||||
Db IleaspMetIleValGlnAsnVal 302
295 |||||

```

RESULT 3
T18629
hypothetical protein B0001.4 - Caenorhabditis elegans
C|Species: Caenorhabditis elegans
C|Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000
C|Accession: T18629
R:Sims, M.
Submitted to the EMBL Data Library, February 1996
A|Reference number: Z18995
A|Accession: T18629
A|Status: preliminary; translated from GB/EMBL/DBJ
A|Molecule type: DNA
A|Residues: 1-229 <MIL>
A|Cross-references: EMBL:Z69634; PTDN:CAA93453.1; GSPDB:GN00022; CESP:B0001.4
A|Experimental source: clone B0001
C|Genetics:
A|Gene: CESP:B0001.4
A|Map position: 4
A|Introns: 21/3; 41/3; 108/2
C|Superfamily: uridine kinase

Alignment Scores:
Pred. No.:      1,03e-32      Length:      229
Score:          512.50       Matches:     105
Percent Similarity: 69.05%   Conservative: 40
Best Local Similarity: 50.00% Mismatches:    58
Query Match:    33.85%      Indels:      7
Db:             2           Gaps:        4

US-09-896-522-3 (1-834) x T18629 (1-229)

QY      CCCTTCGTGATAGGGGTGAGCGCGCACTGCACAGGGGAAGTCGACCGTGTTGAGAAG 126
      ||||| |||||||:::|||||:::|||||:::|||||:::|||||:::|||||::|||::|
Db       8 ProleulinileglyValAladielylthrserCygilylyserThrilevalgiuaag 27

127 ATCATGAGATTGCTGGGACAGAAACAGGTGTGAACAG--CGCACGGAAGGTGTGTCATC 183

```

```
Db 28 ILLEGLUASNLEUASNALEAENALALYGLINSECTLYARGINILEASPILEVALHIS 47
QY 184 CTGAGCCAGGACAGGTTCTCAAGAGTCTTGAACGGCAGACGAAAGCCAAAGCCTTGAAA 243
Db 48 LEUSERTLEUHSERPERHYRARGULLEUSERTLEUVALGULYSLILEUVALAARGUL 67
QY 244 GACAGTACAAATTTTGCATCCAGATCCCTTGATATGATATGATGACAGAGACTCTG 303
Db 68 GLYUPEHESNHEASPHISPROASPGINILEASNHEASPEUEUALAGULTHLEU 87
QY 304 AAGAACATCTGGAGGGCAAAACGTTGAGAGTGGCCGACTATGATTTTGTGACACACTCA 363
Db 88 GLASMERILEASPGILYSTRVALGULILEPROLYSTRYASPMETILETHR---SER 106
QY 364 AGGTTACAGAGACACGAGTGCTTACCCCTGGACGCTGTTCTGTTTGGGCGCATTTG 423
Db 107 SERMEASNGILYTRVALTHRVALGLUPROALALYSLILEILEGLIGULILEU 126
QY 424 GTGCTTACAGCAGGAGATCCGGACATGTTCCACTGGCCTTTCGTTGACACCGAC 483
Db 127 LEUENTYRSPGLUARGVALARGLYSLEUUSERTHYRSLUEUPHEVALGLULYSASN 146
QY 484 TCCGACGTCAGGCTGTCTCGAAGATT-----CTCCGGACGTCGCGCGAGGAGG 534
Db 147 ALAGLUSERARGLEUARGENARGLEUALATHRTYRILEARGSPRYRHISARGALAPRO 166
QY 535 GACCTGAGCAGATTCTGACGCGATACACACTTCTGTAAGCGGCTTCGAGAGATTTC 594
Db 167 -----LEUSERTLEUHSERPERHYRARGULLEUSERTLEUVALAARGULPHE 184
QY 595 TGCCTGCGCAAGAAGATTGCGAGTGTATCATCCACAGAGAGTGCACATATGCTT 654
Db 185 CYSARGPROTHRYSLYSTYRALAASPVALLILELEPROARGLYALAASPHASHSVAL 204
QY 655 GCCATCACTGATCTGTCGACATCCAG 684
Db 205 ALATHRASPLEULEALALYSAENLEUGLN 214
```

RESULT 4

C83809
uridine kinase udk [imported] - *Bacillus halodurans* (strain C-125)
C/Species: *Bacillus halodurans*
C/Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C/Accession: C83809
R/Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A/Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and
A/Reference number: A83650; PMID:20512582; PMID:11058132
A/Accession: C83809
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-211 <STO>
A/Cross-references: GB:AP001511; GB:BA000004; NID:G10173727; PIDN:BAB04994.1; GSPDB:GN00
A/Experimental source: strain C-125
C/Genetics:
A/Gene: udk
C/Superfamily: uridine kinase

Alignment Scores:
Pred. No.: 1,38e-29 Length: 211
Score: 473.00 Matches: 93
Percent Similarity: 66.51% Conservative: 50
Best Local Similarity: 43.26% Mismatches: 56
Query Match: 31.24% Indels: 16
Gaps: 4
DB: 2

US-09-896-522-3 (1-834) x C83809 (1-211)

```
QY 61 CAGCCGCGCTTCTGATAGGGGTGAGCGCGGCACTGCGCAGCGGAAAGTCGACCTGTGT 120
Db 3 LYSARGPROTHRYSLYSTYRALAAGLYGLYTHRCILYSEGLYSTRYTHRVALALA 22
```

```
QY 121 GAGAGATCATGAGTGTCTGGACAGAACGAGTGAACAGCGGACGAGGAGTGTG 180
Db 23 LYSGLILIEPHE-----TYRGLNPHASNGILYSERTILEVAL 35
QY 181 ATCTTAGGCCAGGACAGGTTCTACAAG-----GTCTGACGGCAGGACGAAAGCC 231
Db 36 LEUILEGLIGLINSRPLATYRTRYLSASRGINSERGLNLEUSERTLEUGLILARGLEU 55
QY 232 MAGGCTTGAAGAGACAGTACAAATTTTGCATCCAGATGCTTGTATATGATTTGAG 291
Db 56 -----GLNTHRASNYRASPNSIPROLEUALAPHEASPHASNPHEU 70
QY 292 CACAGGACTGTGAAGACATCTGAGGCGCAAAACGTTGAGAGTGGCCGACTATGATTTT 351
Db 71 ILEGLIHISLEUHSERLEUENASNGILYNALALEGLULYSPROVALTYRASPRTYR 90
QY 352 GTGACACACTCAAGTTACAGAGACACGAGTGCTTACCTTGGACGCTGTTCTGTTT 411
Db 91 LYSLEUHSERTHRARGESRNSGLUVALILEUVALGLUPROLYSAPRYRHISLEU 110
QY 412 GAGGCGATCTTGTGTTCTTACAGCCAGAGATCCGGACATGTTCCACTGGCCTTTC 471
Db 111 GLUGLYILELEUENLEUGLUSPGLUARGLEUARGULTEUENASPRILEYSLIEU 130
QY 472 GTGACACAGGACTCCGACGTCGAGCTGTCTCGAAGATTCTCCGGACGTCGCGC---CGA 528
Db 131 VALASPTHRASPLAASPRILEARGILELEARGARGMEVALARGSPRILEARGULARG 150
QY 529 GGGAGGACCTGGAGCAGATTCTGACGAGTACACACACTTGTGTAAGCGGCTTCGAG 588
Db 151 GLYARGTHREUGLUSERVALILEGLUGINTYRTHRYVALVALARGPROMETHISMET 170
QY 589 GAGTTTGGCTTCCGCAAGAAGATGTCGAGTGTATATCCACGAGAGATGACAT 648
Db 171 GLNPHIELEGLUPROTHRYSLYSTYRALAASPVALLILELEPROGLIGLILYGINASN 190
QY 649 ARGGTTGCATCACTGATCCGTCGACGACATCCAGGACATTCGTG 693
Db 191 ARGVALALALEASPLEUENETVALTHRYSLILEARGALALEILE 205
```

RESULT 5

T41020
probable Uridine kinase - fission yeast (*Schizosaccharomyces pombe*)
C/Species: *Schizosaccharomyces pombe*
C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C/Accession: T41020
R/Murphy, L.; Harris, D.; Wood, V.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, June 1998
A/Reference number: Z21965
A/Accession: T41020
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-454 <MUR>
A/Cross-references: EMBL:AL023860; PIDN:CA119591.1; GSPDB:GN00068; SPDB:SPCC162.11c
A/Experimental source: strain 97Zh-; cosmid c162
C/Genetics:
A/Gene: SPDB:SPCC162.11c
A/Map position: 3

Alignment Scores:
Pred. No.: 2,02e-28 Length: 454
Score: 459.00 Matches: 92
Percent Similarity: 62.33% Conservative: 47
Best Local Similarity: 41.26% Mismatches: 76
Query Match: 30.32% Indels: 8
Gaps: 2
DB: 2

US-09-896-522-3 (1-834) x T41020 (1-454)

```
QY 31 AGCCCGCGCGGAGCGGACGTCGCGACAGCGGCGCTTCTGATAGGGGTGAGCGGC 90
Db 9 SERSERASNPOTHRTRYGLUPROTHRYSPARGLYSVALARGPHEILEGLYLEAGLY 28
```

```

QY 91 GGACCTCCAGCGGAGATGACCGTGTGAGAGATTCATGGAGTTCTGGGACAGAAC 150
Db 29 ProserglyserglytrmservalalaglmeulevallysalaleamleuPro 48
QY 151 GAGGTGAGACGCGGACGCGAGAGGATCTCTGAGCCGACAGGTTCTACAGGTC 210
Db 49 His-----ValValIleleuserleuabserperthyrylyser 61
QY 211 CTGACGCGACAGCAGAGGCCAGGCTTGAAAGACATTCATTTTGACATCCAGAT 270
Db 62 LeuamlaagluglnlyblyaralaphamamabprtyrarpheasrperProglu 81
QY 271 GCCCTTATATATGATTGATGACACAGGACTCTGAAGACATCGTGAGGCGAAACGGTG 330
Db 82 AlalleasprtrphleuLeuPhevallyleuLeuuglnleuysglnlyarglyval 101
QY 331 GAGGTCCGACCTATGATTTTGTGACACACTCAAGGTTACAGAGACACGGTGTCTAC 390
Db 102 AsplleProiletyrserPheamnglwhlasmnargleuProgluThrThrLeuPhe 121
QY 391 CTTGCGGACGCTGCTCTTTGAGGACATCTGTGTTCTACAGCCAGAGATCCGGAC 450
Db 122 GlyAlaserllelelleleuuglnlyllePhehalaleuTyraaprglnlylearyser 141
QY 451 ATGTTCCACCTGCTCTTGTGACACCGACTCCGACGCTGCTGTCTGGAAGATT 510
Db 142 LeuLeuaspvalserValPheleuabprThrapsrseraspvalCysleaserHargyleu 161
QY 511 CTCGGGACCTGCGC---CGAGGAGGAGCCTGAGAGATCTTGACGACGATACACACC 567
Db 162 AsnargrasplleamlytyrargglyArasplleValglyalleuuglnlytyrAsnly 181
QY 568 TTGCTGAACCGCGCTTCGAGGAGATTCTGCTGCCGCAAGAGTATGCCAGATGATC 627
Db 182 PhevallylserProserTyrgluasnphevalargarglnleuserTyrrhrasplelle 201
QY 628 ATCCACGAGAGGACATATATGTTGCCATCACTGATCGAGACATCCAGACAGAC 687
Db 202 ValProargglyaraghaapshlyleuhalalleasphetValIleamPheillearyArg 221
QY 688 ATTCTGAAT 696
Db 222 Thrleuser 224

```

RESULT 6

uridine kinase (EC 2.7.1.48) [imported] - Lactococcus lactis subsp. lactis (strain IL140 D86832

C/Species: Lactococcus lactis subsp. lactis

C/Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001

C/Accession: D86832

R/Bolotin, A.; Winkler, P.; Mauger, S.; Jallion, O.; Malame, K.; Weissenbach, J.; Ehrlich Genome Res. 11, 731-753, 2001

A/Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis ssp A/Reference number: A86625; MUID:21235186; PMID:11337471

A/Accession: D86832

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-206 <STO>

A/Cross-references: GB:AE005176; PID:912724672; PIDN:AAK05758.1; GSPDB:GN00146

A/Experimental source: strain IL1403

C/Genetics:

A:Gene: udk

C/Superfamily: uridine kinase

C/Keywords: phosphotransferase

Alignment Scores:

Pred. No.:	4,43e-28	Length:	206
Score:	454.00	Matches:	92
Percent Similarity:	62.50%	Conservative:	43
Best Local Similarity:	42.59%	Mismatches:	65
Query Match:	29.99%	Indels:	16
DB:	2	Gaps:	5

```

US-09-896-522-3 (1-834) x D86832 (1-206)
QY 61 CAGCGCCCTTCTGATTAAGGGTGAAGCGGCGGACCTGCCAGCGGAAATCGACCTGTGT 120
Db 2 LysIleThrleuIlelleleuglyalThrnglyserlaaserCdylyserValSer 21
QY 121 GAGACATGATGAGGTGCTGGGACAGACAGAGTGAACAGCGGACGCGAAGTGTGTC 180
Db 22 HisAlIleleuIlnhrhe---Seramngl-----ArgIleAla 34
QY 181 ATCCTGAGCCAGACAGGTTCTTACAG-----GTCTGACGCGACAGACAAAGCC 231
Db 35 MetIleGlwhlIeasrertyrtyrlyasrgrlnserHlleuThrPhegluglnArgthr 54
QY 232 AAGGCTTGAAGAGACAGTACATTTTGACSCATTCGAGAGGCTTGATATGATTGAG 291
Db 55 LysThr-----AsnytyraapshsProleuAlaPheasprhraspyrlyeu 69
QY 292 CACAGACTCTGAAAGAACATCGTGAGGCGCAACCGTGAGAGTCCGACTATGATTTT 351
Db 70 IleAlglInleuylseleuenglnlyrgrlyaralaleValasplleProiletyrAspyr 89
QY 352 GTACACACTCAAGTTACCAAGACACACGAGTGTCTACCTGCGGACGCTGTCTGTT 411
Db 90 AlaIyHlelthrxargserglnlythrtyrArgglnlProvalasPvalleuIleVal 109
QY 412 GAGGGATCTGTGTTCTTACAGCCAGAGATCCGGAGACATGCCAGGCTCTTC 471
Db 110 GlnglylleuvalleuuglnleuabprGluArgleuArgAspleuIeasrlyleuIlePhe 129
QY 472 GTGACACGAGCTCCGACGTCAGGCTGTCTCGAAGATTCTCCGGACGCTG---CGCGA 528
Db 130 ValasprThrapsraspvalargllelleargarglleargargrsplleuglnlyArg 149
QY 529 GGGAGGACCTGAGACAGATTCTGACGCGATACACCACTTGTGAGCCGCGCTTCGAG 588
Db 150 GlyArgThrleuabserAllellethrglnlyrleuaspAlaVallyserProweTyrrHs 169
QY 589 GATTTGCTGCTCGCGCAAGAGTATGCCAGTGTATATCCACGAGGAGTGAACAAT 648
Db 170 GlnPheIleuglnProthrlyasrtyrAlaasPvalIleleProgluglyalSeram 189
QY 649 ATGTTGCCATCAACTGATCTGCGACACATCCAGACATCTTGAAT 696
Db 190 ThrValGlyValaspllellethrThrlylleAlaserlleuasn 205

```

RESULT 7

uridine kinase [imported] - Staphylococcus aureus (strain N315)

C/Species: Staphylococcus aureus

C/Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001

C/Accession: B89943

R/Kuroda, M.; Ohba, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguchi, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.; Shiba, T.; Hatcori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K. Lancet 357, 1225-1240, 2001

A/Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus. A/Reference number: A89758; MUID:21311952; PMID:1143146

A/Accession: B89943

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-207 <KUR>

A/Cross-references: GB:BA000018; PID:g13701409; PIDN:BA842703.1; GSPDB:GN00149

A/Experimental source: strain N315

C/Genetics:

A:Gene: udk

C/Superfamily: uridine kinase

Alignment Scores:

Pred. No.:	1.59e-27	Length:	207
Score:	447.00	Matches:	88
Percent Similarity:	65.09%	Conservative:	50
Best Local Similarity:	41.51%	Mismatches:	58

[illegible]

Pred. No.:	8	28e-27		Length:	212
Score:	438.00		Matches:	88	
Percent Similarity:	60.75%		Conservative:	42	
Best Local Similarity:	41.12%		Mismatches:	68	
Query Match:	28.93%		Indels:	16	
DB:	2	Gaps:	4		

US-09-g96-e522-3 (1-834) x B95140 (1-212)

```

QY      CGGCGCTTCGTATAGGGTGAAGCGGCGGCACGTGCCAGCGGAAATTCCACCCTGTGTGAG   123
        ||||| ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db     AArgProIleIleIleGIValInnIGlySeNGlYglYglYstHrSerValSerArg   23
QY      124 AAGCATGTGAAGTTTGCTGGACAAGAAGCAGAGTGGAAACACGCGAAGCCGGAAGTGTGCATC   183
        |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db     24 AlaIleLeuSer-----HisPheProAspGluYSilSeMet    36
QY      184 CTGAGCGCAAGCACAGTTCTTAACAAG-----GTCTGACGGCGACAGAACGCCCAAG   234
        :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db     37 IIGlnhiIsasPseryTYrYLysAbpGLnsenHiSiSleuthrHeGluDAUGValLYs   56
QY      235 GCGTTGAAAAGACAGTCAATTTTGACSCATCCAGATGCGCTTGATAATGATTGATGAC   294
        |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db     57 Thr-----AsnTYraSPHisPRORheAlahesprHtraBleMeIle   71
QY      295 AGGACTGTGAAGAACATCGTGGAGGCGAAACSGTGGAGGTGCCACTATGATTTTGTG   354
        ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db     72 GluGINileLysGIuleuLeuAlaGlAYarProValAspILlePrOfHrTYraerTYthr   91
QY      355 ACACATCTCAAGGTtaccAgagAcccAGctgStgtCAccTtcGGAcCGtgTtCGttTAG   414
        ||::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db     92 GluhISThrXargSerSerLystrHYrarGlnGlnuProDinspValPheIIeValGlu   111
QY      415 GGCACTTGTGGNTCTTACAGCCAGAGATCCGGGACATGTTTCCACCTGCGCCTTTCGTG   474
        |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db     112 GIyIlleuValILeuGIAspRlySArgLEuaYqASpIeumetaspRIelySIIPheVal   131
QY      475 GACACGCACTCCGACGTCAAGGCTGTCTCGAAGATTTCCGGAGACGTT---CGCGAGGG   531
        |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db     132 AsphrrAspArspArValARgIlelleArxArgILEyarGrAmpeMetGIuIdArgGIy   151
QY      532 AGGCACTCGAGCAGATTCTGACGCGACGTACACCACTTCTGTGAAGCCGCGCTTCGAGAG   591
        |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db     152 ArgSerIeuAspservallIElaemGIntYLEugIyValIalYsprometyrhNISgLn   171
QY      592 TTTCGCGCTCGCGAAAGAGTAGTGCAGATGATATATCCCAGAGGAGTGACATATG   651
        |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db     172 PheIlegIuseThrLYsaRGTYrALaAspRIlevalIleProGUguIyaLSeraenthr   191
QY      652 GTTGCATCAACCTGATCGTGCGACCATCCAGACATTTCTG   693
        |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db     192 ValAlaIleAspIeuLeuthrTHrySlleAlalyaleIeu   205

```

RESULT 9

A99008
uridine kinase (EC 2.7.1.48) [imported] - Streptococcus pneumoniae (strain R6)
CSpecies: Streptococcus pneumoniae
CDate: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 02-Nov-2001
CAccession: A99008
E:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; Dehoff, B.S.; E.
F.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.O.; Lu, J.; Matsushima, P.; McAhren, S.; M
J., P.; Sun, P.M.; Winkler, M.E.
Y.: Bacteriol. 183, 5709-5717, 2001
A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;
A>Title: Genome of the Bacterium Streptococcus pneumoniae Strain Re.
A:Reference number: A97872; PMID:21429245; PMID:11544234
A:Accession: A99008
A>Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-212 <KIR>
A:Cross-references: GB:AEO07317; PIDN:AAK9893.1; PID:g15458714; GSPDB:GN00174
!Genetics:
!Gene: udk

Dh 185 ValAspAlaAspLeuAspValCysLeuAlaArgHrGluSerArgAspAlaValSerArg 204
Qy 529 GGGAGGAGGAGCTGGAGGAGATTCTGAGCGAGTACCACTTCGGAAGCGGCGCTTCGAG 588
Db 205 G1AArgAspLeuAspGlyCysIleGlnGlnTrpGlnLysPheValLysProAsnAlaVal 224
Qy 589 GAGTTTCGCTGCGCCGCAAAAGATATCCGATGTCATCCACAGAGAGTGACAAAT 648
Db 225 LysPheValLysProThrMetLysAsnAlaAspAlaIleLeuProSerMetSerAspAsn 244
Qy 649 ATGATTCATCAACCTGATCGTGCAGCAGCATC----- 681
Db 245 AlAThrAlaValAsnLeuIleIleAsnHisIleLysSerLysLeuGluLeuLysSerAsn 264
Qy 682 -----CAGCATTCGGAAGGT 699
Db 265 GluHisLeuArgGluLeuIleLysLeuGlySerSerProSerGlnAspValLeuAsnArg 284
Qy 700 GACATC 705
Db 285 AsnIle 286

RESULT 11
G69728
uridine kinase udk - Bacillus subtilis
C/Species: Bacillus subtilis
C/Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 21-Jul-2000
C/Accession: G69728
R/Kunze, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berton
C.; Bron, S.; Brouillet, S.; Buschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chd
A.; Ehrlich, S.D.; Emmerison, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A/Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallier
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holtsappel, S.; Hosono, S.; Hullo, M.F.
Koelter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Landinote,
A.; Luthers, Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelli
Rieger, M.; Rivolta, C.; Roche, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon,
A.; Luthers, Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpestra, P.; Tognoni, A.; Toesio, V.; Uchiyama,
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
A/Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.
A/Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A/Reference number: A69580; MUID:98044033; PMID:9384377
A/Accession: G69728
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-211 <GUN>
A/Cross-references: GB:299117; GB:AL009126; NID:g2634966; PIDN:CAB14675.1; PID:g2635179
A/Experimental source: strain 168
C/Genetics:
A/Gene: udk
C/Superfamily: uridine kinase

Alignment Scores:
Pred. No.: 3,57e-26 Length: 211
Score: 430.00 Matches: 84
Percent Similarity: 63.26% Conservative: 52
Best Local Similarity: 39.07% Mismatches: 63
Query Match: 28.40% Indels: 16
Gaps: 4
DB: 2

US-09-896-522-3 (1-834) x G69728 (1-211)

Qy 61 CAGCGGCGCTTCCTGATAGGGGTGAGCGGCGGCACTGCAGCGGGAAGTGCACCTGTGT 120
Db 3 LysAsnProValValIleGlyIleAlaGlyIleAlaGlySerGlySerValThr 22
Qy 121 GAGAAGATCATGAGTTCGCTGCGAGACAGAGTGAACAAGCGGCGGCAAGAGTGTGTC 180
Db 23 ArgSerLysIleArg-----GlnPheLysGlyHisSerIleLeu 35
Qy 181 ATCCTGAGCGAGACAGGTTCTACAG-----GTCTGAGCGGCGAGCAAGAGGCC 231

Db 36 MetIleGlnGlnAspLeuTrpLysAspGlnSerHisLeuProPheGlnGluArgLeu 55
Qy 232 AAGCGCTTGAAGAGACAGTACAAATTTTGAACCATCCAGATGCCCTTTATATGATTTGAAG 291
Db 56 AsnThr-----AsnTrpAspHisProLeuAlaPheAspAsnSerTrpLeu 70
Qy 292 CACAGAGACTGTGAGAAACATCGTGAGGGGCAAAAGCGTGAAGGTGCGACCTTATGATTTT 351
Db 71 IleGlnHisIleGlnAspLeuAsnTrpArgProIleGlnLysProIleTrpAspTrp 90
Qy 352 GTGACACACTCAAGTTTACCGAGACACAGCGTGTCTTACCTTCGCGAGCGTGTCTT 411
Db 91 LysLeuHisThrArgSerGlnGlnThrValHisValGlnProLysAspValIleIleLeu 110
Qy 412 GAGCGCATCTGTGTTCTTACAGCCAGAGATCCGAGACATGTTCCACCTGCGCTTTC 471
Db 111 GluGlyIleLeuValLeuGlnAspLysArgLysAspLeuSerAspIleLysLeuTrp 130
Qy 472 GTGACACAGCATCGGAGCGTGTCTTCCAGAGGTTCTCCGAGCGTGTCTTCAG 528
Db 131 ValAspThrAspAlaAspLeuArgIleIleArgArgIleMetArgAspIleAsnGluArg 150
Qy 529 GGGAGGAGGAGCTGGAGGAGATTCTGAGCGAGTACACCACTTCGTCGAGCGGCTTCGAG 588
Db 151 GlyArgSerIleAspSerValIleGlnGlnTrpValSerValValArgPromethisAsn 170
Qy 589 GAGTTTCGCTGCGCCGCAAAAGATATCCGATGTCATCCACAGAGAGTGACAAAT 648
Db 171 GlnPheValGlnProThrLysArgTrpAlaAspIleIleProGlnGlyGlnAsn 190
Qy 649 ATGATTCATCAACCTGATCGTGCAGCAGCATC----- 693
Db 191 HisValAlaIleAspLeuMetValThrLysIleGlnThrIleLeu 205

RESULT 12
F84277
uridine kinase (imported) - Halobacterium sp. NRC-1
C/Species: Halobacterium sp. NRC-1
C/Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 17-May-2002
A/Accession: F84277
R/M, W.V.; Kennedy, S.P.; Mahafas, G.G.; Berguist, B.; Pan, M.; Shukla, H.D.; Lasky, S.
; Leithauer, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablor
Jung, K.H.; Alam, M.; Freltas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A/Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ehardt, H.; Lowe, T.M.; Lie
A/Title: Genome sequence of Halobacterium species NRC-1.
A/Reference number: A84160; MUID:20504483; PMID:11016950
A/Accession: F84277
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-248 <STO>
A/Cross-references: GB:AE004437; NID:gl0580750; PIDN:AGI19586.1; GSPDB:GN00138
C/Genetics:
A/Gene: urk
C/Superfamily: uridine kinase

Alignment Scores:
Pred. No.: 3.94e-25 Length: 248
Score: 417.00 Matches: 83
Percent Similarity: 63.64% Conservative: 50
Best Local Similarity: 39.71% Mismatches: 58
Query Match: 27.54% Indels: 18
Gaps: 5
DB: 2

US-09-896-522-3 (1-834) x F84277 (1-248)

Qy 70 TTCCTATAGGGGTGAGCGGCGGCACTGCAGCGGGAAGTGCACCTGTGTGAGAGATC 129
Db 29 PheAlaIleGlyIleLeuAlaGlyLysTrpAlaGlyLysTrpValAlaArgGluIle 48
Qy 130 ATGAGTTCTGCGAGACAGAGTGTGAACAGCGGCGGCAAGAGTGTCTCATCTCGAGC 189

Dh 49 ThrAspAnValGlyIuSer-----AlaThrLeuIlePro 60
Qy 190 CAGAGAGGTTCTCAAGGTCCTGACG-----GCAGAGCAGAAAGGCCAAGGCTTGG 240
Db 61 LeuAspAnValTyGlyIuSerLeuAspArgProPheGluGluArgAlaAsnIle--- 79
Qy 241 AAGAGCAGTCAATTTTGAACCATCCAGATGCTTTGATATGATTTGATGACAGAGACT 300
Db 80 -----AanTyAspHisProSerAlaPheGluTyrGluLeuLeuArgThrHis 95
Qy 301 CTGAAGAACATCTGCGAGAGGCAAAACGATGAGTCCGACCTTATGATTTTGTACACAC 360
Db 96 IleAspAlaLeuSerGlyGlnSerIleGluMetProGlnTyAspPheGluArgHis 115
Qy 361 TCAAGGTTACCGAGACACCGAGTGTCTACCTCGAGCGGCGTTCGTTGAGGGATC 420
Db 116 ValArgLysAlaAspArgValValValGluProThrAspValIleValLeuGluGlyIle 135
Qy 421 TTGGTGTCTTACACCGAGAGATCCGGACATGTTCCACCTGCGCTCTTCGTGAGACAC 480
Db 136 LeuAlaLeuSerAspGluThrValAlaAspMetLeuAspLeuHisIleTyValGluThr 155
Qy 481 GACTCCGACGTCAAGCTGTCTCGAAGACTTCCGGGAC--GTCCGCGAGGAGGAGAC 537
Db 156 AspAlaAspValArgIleLeuArgArgIleGluArgAspValValGluArgGlyArgGlu 175
Qy 538 CTGAGAGAGATTCTGAGCAGATACACACCTTCGTGAAGCGCGCTTCGAGGAGTTCTGC 597
Db 176 LeuGluGlyValMetAspGlnTyLeuSerThrValIlyAspPheMetHisGluGlnPheIle 195
Qy 598 CTGCGCAAGAAAGATGCGCATGTGATCATCCACGAGGATGACAAATATGTTGCC 657
Db 196 GluProThrArgLysArgHisAlaAspIleIleIleProGluGlyAla---AanSerValAla 214
Qy 658 ATCAACCTGATCGTCGACGACATCCAG 684
Db 215 ValAsnLeuSerGlnGluLysValGln 223

RESULT 13

D86582
uridine kinase [imported] - Chlamydia pneumoniae (strain J138)
C/Species: Chlamydia pneumoniae, Chlamydia pneumoniae
C/Date: 02-Mar-2001 #sequence revision 02-Mar-2001 #text_change 23-Mar-2001
C/Accession: D86582
R/Shirai, M.; Hixkawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Ise
Nucleic Acids Res. 28, 2311-2314, 2000
A/Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.
A/Reference number: A86491; MUID:20330349; PMID:10871362
A/Accession: D86582
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-222 <STO>
A/Cross-references: GB:BA000008; NID:g8979107; PIDN:BAA98942.1; GSPDB:GN00142
A/Experimental source: strain J138
C/Genetics:
A/Gene: CP10735
C/Superfamily: uridine kinase

Alignment Scores:
Pred. No.: 9.65e-25 Length: 222
Score: 412.00 Matches: 84
Percent Similarity: 63.01% Conservative: 54
Best Local Similarity: 38.36% Mismatches: 59
Query Match: 27.21% Indels: 22
DB: 2 Gaps: 6

US-09-896-522-3 (1-834) x D86582 (1-222)

Qy 73 CTGATAGGAGTGAAGCGGCGGAGTCAACCGTGTGAGAGATCATG 132
Db 8 IleIleGlyIleThrGlyGlySerGlyAlaGlyLysThrThrLeuThrGlnAsnIleLys 27
Qy 133 GAGTTCTGGAGACAGAAAGAGTGAACAGCGGACGGAAGTGTATCTGAGCCAG 192

Dh 28 GluIlePheGlyGluAsp-----ValSerValIleGlyGln 39
Qy 193 GACAGGTTCTACAG-----GTCTGACGCGACAGACAGAAAGCCCAAGGCTTGAAA 243
Db 40 AspAsnTyTyTyIleAspArgSerHisTyTyTyProGluGluArgAlaAsnLeuIle--- 58
Qy 244 GACAGTACATTTTGAACCATCCAGATGCTTTGATATGATTTGATGACAGAGCTTG 303
Db 59 -----TrpAspHisProSerAlaPheAspAspLeuLeuIleSerAspIle 74
Qy 304 AAGACATCTGAGAGGCAAAACGATGAGTCCGACCTTATGATTTTGTACACACTCA 363
Db 75 LysArgLeuLysAsnAsnGlnIleValGlnAlaProValPheAspPheValLeuGlyAsn 94
Qy 364 AGG---TTACCAAGACACACGAGTGTCTACCTCGGACGAGTGTCTTTGAGGGCATC 420
Db 95 ArgSerLysThrIleGluThrIleGluThrIleTyProSerLysValIleLeuValGluGlyIle 114
Qy 421 TTGGTGTCTTACACCGAGAGATCCGGACATGTTCCACCTGCGCTCTTCGTGAGACAC 480
Db 115 LeuValPheGluAsnGlnGlnLeuAlaGlyAspLeuMetAspIleArgIlePheValAspThr 134
Qy 481 GACTCCGACGTCAAGCTGTCTCGAAGACTTCCGGGACGCG--CGAGGAGGAGAC 537
Db 135 AspAlaAspGluArgIleLeuArgArgMetValArgAspValGlnGluGlyAspSer 154
Qy 538 CTGAGCAGATTCTGACGACGATCACACACCTTCGTGAAGCGCGCTTCGAGGAGTTCTGC 597
Db 155 ValAspCysIleMetSerArgTyLeuSerMetValLysProMetHisGluLysPheIle 174
Qy 598 CTGCGCAAGAAAGATGCGCATGTGATCATCCACGAGGATGACAAATATGTTGCC 657
Db 175 GluProThrArgLysTyTyAlaAspIleIleValHisGlyAsnTyArgGlnAsnValVal 194
Qy 658 ATCAACCTGATCGTCGAG-----CACATCCAGACATTTCTGAATGTGAC 702
Db 195 ThrAsnIleuSerGlnLysIleLysAsnHisLeuGluAlaAsnAlaLeuGluSerAsp 213

RESULT 14

E72041
uridine kinase CP0011 [imported] - Chlamydia pneumoniae (strains CWL029 and AR39)
C/Species: Chlamydia pneumoniae, Chlamydia pneumoniae
C/Date: 23-Apr-1999 #sequence revision 23-Apr-1999 #text_change 11-May-2000
C/Accession: E72041; E81622
R/Kalman, S.; Mitchell, W.; Marathe, R.; Lemmel, C.; Fan, J.; Olinger, L.; Grimwood, J.;
Nature Genet. 21, 385-389, 1999
A/Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.
A/Reference number: A72000; MUID:99206060; PMID:10192388
A/Accession: E72041
A/Molecule type: DNA
A/Residues: 1222 <ARN>
A/Cross-references: GB:AE001655; GB:AE001363; NID:g4377039; PIDN:AAD18874.1; PID:g4377041
A/Experimental source: strain CWL029
R/Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heideberg, J.F.; White, O.; Hickey, J.
C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,
Nucleic Acids Res. 28, 1397-1406, 2000
A/Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
A/Reference number: AB1500; MUID:20150255; PMID:10684935
A/Accession: E81622
A/Molecule type: DNA
A/Residues: 1-222 <REA>
A/Cross-references: GB:AE002165; GB:AE002161; NID:g7188948; PIDN:AAF37907.1; PID:g7188950
A/Experimental source: strain AR39, HL cells
C/Genetics:
A/Gene: CPn0735; CPn011
C/Superfamily: uridine kinase

Alignment Scores:
Pred. No.: 9.65e-25 Length: 222
Score: 412.00 Matches: 84
Percent Similarity: 63.01% Conservative: 54
Best Local Similarity: 38.36% Mismatches: 59

```

Query Match:      27.21%      Index:      22
DB:               2          Gaps:      6

US-09-896-522-3 (1-834) x E72041 (1-222)

QY      73 CTGATAGGGGTGAGCGCGGCACTGCGAGCGGAGATGCACTGTGTGAGAAATCATG 132
      8 11e11e61y11e1hrrg1y1serg1y1a1g1y1shrr1h1euthr1h1a1n1le1y1e 27
QY      133 GAATTGCTGGGACAGAACGAGTGGAAACAGCGGACGGAGGATGTCATCTGAGCCAG 192
      28 G1u11ePheg1y1g1u1a1sP-----ValSerVal11e1Cy5g1n 39
Db      193 GACAGGTTCTACAG-----GTCTGACGGACAGAGAGAGAGGCGCAAGCCTTGAAA 243
      40 AspAsm1y1r1y1y1u1y1a1sP1a1r1g1Ser1h1s1Y1r1h1Pro1u1u1a1a1n1le1u11e1-- 58
QY      244 GACAGTACATTTTGACATCCAGATGCTTGTATGATGATTTGATGACAGAGACTCTG 303
      59 -----T1rPasp1h1sP1roA1sP1a1a1Phe1a1sP1a1sP1a1n1le1u11e1Ser1a1P1le 74
QY      304 AAGAATCTCGTGAAGGCAAAACGGTGGAGTGGCCGCACTATGATTTTGTGACACACTCA 363
      75 1y1a1r1g1u1e1y1a1sP1a1n1s1n1u11e1Val1G1n1a1P1ro1Val1Phe1a1sP1he1Val11e1u1g1y1a1sP1n 94
QY      364 AGG---TTACACAGACACACAGCGTGTCTACCGTGGTGTCTGAGGAGGATC 420
      95 A1r1g1Ser1y1shrr1g1u11e1g1u1hrr11e1y1r1P1ro1Ser1y1s1Val11e1u1e1Val1G1u1g1y11e 114
Db      421 TTGGTGTCTTACAGCCAGAGATCCGGAGATGTTCCACCTGCGCCTTCTGTGACACAC 480
      115 11e1Val1Phe1g1u1a1n1s1n1g1u11e1u1a1r1g1a1sP1le1u1a1r1g11e1Phe1Val1a1P1hrr 134
QY      481 GACTCCGACSTCGAGCTGTCTCGAAGATTTCTCCGGACSTGCGC--CGAGGAGGAGAC 537
      135 Asp1a1a1sP1a1r1y1a1r1g11e1u1a1r1g1Met1Val1a1r1g1y1a11G1n1g1u1g1y1a1sP1Ser 154
QY      538 CTGGACAGATTTCTGACGCGAGTACACACCTTGATGAGCGGCGCTTCGAGAGATTTCTGC 597
      155 Val1a1sP1ro1y11e1Met1Ser1a1r1y1r1e1u1Ser1Met1Val1y1e1P1ro1Met1h1s1G1u1y1Phe11e 174
Db      598 CTCCCGACAAAGAGATGCGCGATGTCATGATCCACAGAGATGACATATGCTTGCC 657
      175 G1u1P1ro1hrr1a1r1g1y1r1a1a1sP11e11e1Val1h1e1s1y1a1n1y1r1a1r1G1n1a1n1Val1 194
QY      658 ATGACCTGATCGTGAC-----CACATCCAGACATTTCTGAATGGTGAC 702
      195 Thr1a1n11e1u1e1u1Ser1g1n1y11e1y1a1sP1h1s11e1u1G1n1a11e1u1G1n1Ser1a1sP 213

RESULT 15
G70101
uridine kinase (udk) homolog - Lyme disease spirochete
C/Spectes: Borrelia burgdorferi (Lyme disease spirochete)
C/Pate: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 17-Nov-2000
C/Accession: G70101; T46979
R/Praeer: C.M.; Caslens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White
son, D.; Peterson, J.; Kerlavage, A.R.; Quackendush, S.; Salzberg, S.; Hanson, M.; Vugt,
; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
Nature 390, 580-586, 1997
A/Authors: Smith, H.O.; Venter, J.C.
A/Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.
A/Reference number: A70100; PMID:98065943; PMID:9403685
A/Accession: G70101
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-207 <K1E>
A/Cross-references: GB:AE001115; GB:AE000783; NID:g2687879; PIDN:AA66392.1; PID:g268788
A/Experimental source: strain B31
R/Bouraux-Bode, C.; Margarita, D.; Gilles, A.M.; Barzu, O.; Gitons, I.S.
FEMS Microbiol. Lett. 151, 257-261, 1997
A/Title: Borrelia burgdorferi uridine kinase: an enzyme of the pyrimidine salvage pathwa
A/Reference number: Z24339; PMID:97372541; PMID:9228761
A/Accession: T46979

```

```

A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 2-207 <BOU>
A/Cross-references: EMBL:X97449; NID:g1405440; PIDN:CAA66081.1; PID:g1405442
A/Experimental source: strain HB19, ssp. sensu stricto
C/Genetics:
A/Gene: udk
C/Superfamily: uridine kinase

Alignment Scores:
Pred. No.:      1,26e-24      Length:      207
Score:          410.50      Matches:      79
Percent Similarity: 59.81%      Conservative: 49
Best Local Similarity: 36,92%      Mismatches: 59
Query Match:    27.11%      Indels:      27
DB:             2          Gaps:      4

US-09-896-522-3 (1-834) x G70101 (1-207)

QY      73 CTGATAGGGGTGAGCGCGGCACTGCGAGCGGAGATGCACTGTGTGAGAAATCATG 132
      5 11e11e61y11e1e1r1g1y1ser1g1y1s1r1g1y1shrr1h1a1n1le1y1e1Ser 24
Db      133 GAATTGCTGGGACAGAACGAGTGGAAACAGCGGACGGAGGATGTCATCTGAGCCAG 192
      25 G1u1Phe11eP1ro1g1u1-----Phe1Val11e1u1e1Ser1G1n 35
QY      193 GACAGGTTCTACAGGTTCTGACGCGAGACAGAGGCCAAGGCTTGAAAGACACTAC 252
      36 AspAsm1y1r1y1y1u1y1a1sP1a1r1g1Ser1h1s1Y1r1h1Pro1u1u1a1a1n1le1u11e1--G1y1a1sP1y1 45
Db      253 -----A1TTTGA1SCAT1TCCAGAT1GCTTGTATGAT1TTGAT1GATG 291
      46 G1u1h1s1G1u1Phe1Ser1y1s1Val1a1sP1he1a1r1h1sP1roA1a1a1Phe1a1sP1a1sP1a1n1le1u11e1Phe 65
QY      292 CACAGGATCTGTAAGAACATCTGTGAGGCAAAACGGTGGAGTGGCCGCACTATGATTTT 351
      66 1y1r1G1n11e1u1e1y1a1sP1a1n1e1u1y1y1a1sP1ro11e1a1sP1e1r1o11e1u1y1r1a1r1Phe 85
Db      352 GTGACACACTCAAGTTACCAAGACACACGAGTGTCTACCTGCGGACSTGCTGTGTT 411
      86 11e1a1n11e1y1a1r1g1n1e1u1y1shrr1h1e1u1a11P1ro1hrr1P1ro1Val11e1Val1 105
QY      412 GAGGAGATTTGATGTTCTTACAGCCAGAGATCCGGAGATGTTCCACCTGCGCCTTTC 471
      106 G1u1g1y11e1Met11e1Phe1Val1G1u1a1r1g1Val1a1r1g1a1sP11e1a1sP1e1u11e1a1sP1e1y11e1y1 125
Db      472 GTGGACACGAGCTCCGAGCTGCTCGAAGAGTCTCCGGAGCTG--CGCCGA 528
      126 11e1a1r1hrr1P1ro1a1n1sP11e1a1r1g1Phe11e1a1r1g1u1a1r1g1a1r1g1a1sP11e1Ser1y1a1r1g 145
QY      529 GAGAGGAGCTGGAGACAGATTTGACGACAGTACACACCTTGTAAGCGGCGCTTGAG 588
      146 G1y1a1r1hrr1Val1G1u1Ser1Val11e1a1r1G1n1y1r1e1u1a1n1hrr1hrr1a1r1g1r1g1y1r1y1 165
Db      589 GAATTGCTGCTCCGACAAAGATGCGCGATGTCATGATCCACAGAGATGAGAT 648
      166 A1r1g1Phe11e1G1u1P1ro1hrr1y1e1g1u1y1r1a1a1sP11e11e1P1ro1u1G1y1g1h1a1n 185
QY      649 ATGGTTGCATCAACTGATCGTGACGACATTCACAGACATTCAGAGACATT 690
      186 Asp1y1a11e1u1y1r1Val11e1u1e1Ser1hrr1Phe11e1u1y1Ser1e1u 199

Search completed: November 25, 2003, 07:40:43
Job time : 26.0008 secs

```

This Page Blank (uspto)


```
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatchches: 0
Query Match: 95.77% Indels: 0
DB: 9 Gaps: 0
```

US-09-896-522-3 (1-834) x US-09-896-522-2 (1-277)

QY	1	ATGGGCTTCGGACGGAGAGCCGAAAGCTCTCAGAGAGCCCGCGCGGAGGCGCAACGCTCCGAC	60
Db	1	MetAlaSerAlaGlyGlyGlnAspCysGlnSerProAlaProGlnAlaAspArgProHis	20
QY	61	CAGCGGCGCTTCCTGTCTAGAGGGGTGAGCGGCGGCGACTCCAGCGGGAAGTGACCGTGTCT	120
Db	21	GlnArgProPheLeuIleGlyValSerGlyThrAlaSerGlyLysSerThrValCys	40
QY	121	GAGAAAGATATATGAGATTGCTGGGACGAAACAGAGGTGGAACGGCGGACGGGAAGGTGCTC	180
Db	41	GlnLysIleMetGlnLeuLeuGlnIleAsnGlnValGlnGlnArgGlnArgLysValVal	60
QY	181	ATCCGAGGCGCAGACAGAGTTCTTACAAGGTCCTTGACGCGACAGAGAAAGCCAGAGGCTTG	240
Db	61	IleLeuSerGlnAspArgPheTyrLysValLeuThrAlaGlnGlnLysAlaLysAlaLeu	80
QY	241	AAAGACAGTACATTTTGACCATCCAGATGCCCTTGATATGATTGATGACACAGACT	300
Db	81	LysGlyGlnTyrAsnPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThr	100
QY	301	CTGAAGAATATGTTGGAGGGCAAAAGCGTGGAGGTCCGACCTATATTTTGTGACACAC	360
Db	101	LeuLysAsnIleValGlnGlyLysThrValGlnValProThrLysAspPheValThrHis	120
QY	361	TCNAGGTTACACAGAGCCACGCGTGTCTACCCCTGCGGACGCTGTTGTGTTGAGGCGATC	420
Db	121	SerArgLeuProGlnThrThrValValTyrProAlaAspValValLeuPheGlnGlyIle	140
QY	421	TTGGTGTTCTACAGCCAGAGAGATCCGGGACATTTCCACTCGCGCTCTTCGTGACACC	480
Db	141	LeuValPheTyrSerGlnGlnIleValArgAspMetPheHisLeuValArgLeuPheValAspThr	160
QY	481	GACTCCGACGTACAGGCTGTCTGGAAGATTTCTCCGGAAGTGGCGGAGGAGGAGACCTG	540
Db	161	AspSerAspValAlaArgLeuSerArgValIleValArgAspAlaArgArgGlyArgAspLeu	180
QY	541	GAGCAGATTCTGACGCGAGTACACCACTTCGTGAAACCGGCGCTTCGAGAGATTCTCCTG	600
Db	181	GlnGlnIleLeuThrGlnTyrThrThrPheValLysProAlaPheGlnGlnPheCysLeu	200
QY	601	CCGACAAAGAGTATGCCATGTATCATCCACGAGAGGTGGACAATATGTTGCCATC	660
Db	201	ProThrLysLysTyrLysAlaAspValIleIleProArgGlyValAlaAspAsnMetValAlaIle	220
QY	661	AACCTGATGTGCAGACATCCAGAGCATTTGTAAGTGGACATCTGCAAATGGACCGA	720
Db	221	AsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysLysThrPheHis	240
QY	721	GGAGGATCCAAATGGCGGAGCTACAAGCGGACCTTTTCTGAGCCAGAGGACCACTCTGG	780
Db	241	GlyGlySerAsnGlyArgSerTyrLysArgThrPheSerGlnProGlyAspHisProGly	260
QY	781	ATGCTGACCTCTGGCAAAAGGTCACATTTGGAGTCCAGACGACGACCCAC	831
Db	261	MetLeuThrSerGlyLysArgSerHisLeuGlnSerSerSerArgProHis	277

```

RESULT 2
US-09-925-300-1160
; Sequence 1160, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben,
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: P401
; CURRENT APPLICATION NUMBER: US/09/925,300

```

```

CURRICULUM VITAE: 2001-08-10
PRIORITY APPLICATION NUMBER: PCT/US00/05988
PRIORITY FILING DATE: 2000-03-08
PRIORITY APPLICATION NUMBER: 60/124,270
PRIORITY FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 1890
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1160
LENGTH: 337
TYPE: PROT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (38)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (46)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (155)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (169)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-300-1160

Alignment Scores:
Pred. No.:          9,18e-68          Length:          337
Score:              897.50            Matches:         191
Percent Similarity: 73.31%           Conservative:    26
Best Local Similarity: 64.53%         Mismatches:     45
Query Match:        59,288            Indels:         34
OB:                 10                Gaps:           6

```

US-09-896-522-3 (1-834) x US-09-925-300-1160 (1-337)	
QY 3 GGGCTTGGCGGGAGGCGA-----AGACTGGAGAGCCCGCGCGGAGGCCGA---50	
Db 57 GYAlAGlAGlGYlAaRgHlAsErVAlArThRgIySeRgIyARgARgIGlGYlAaSn76	
QY 51 -----CCGCGCCGACCA-----GGGGCC68	
Db 77 HlSGlYARgIYGlAaRgHlAsPRoAlaSPRoAlaGluPRoAlaAlaGlAaRgARgARgAla96	
QY 69 CTTCCTGATAGGGGTGAGCGCGCGGACATCGACGCGGGAGATCGACCTGTGTGAGAAAT128	
Db 97 LeuPRoTYARgARgHlS-GlYgIYThAlAsErClYbYsErSeRAlYbAlAlaYsII116	
QY 129 CATGAGTGTCTGGGACAGAAACAGTGGAAACAGCGGACGCGGAGGTGATCCTGAG188	
Db 116 eValGlnLeuLeuGlgYlGlnAeNgIuValAsPTYARgGlnbYgInValIleuSe136	
QY 189 CGAGGACAGGTTCTTCAGAGTCCCTGACGGGACAGACAGAAAGGCCCTTGAAGAGCA248	
Db 136 RgInAsPeRPeTYARgVAlLeuThSeRGlGlnYAlaYsAlaLeuYs**G1156	
QY 249 GTACATTTTGGACCAATCCAGATGCTTTGATTAATGATTTGTCACACGACTCGAAGAA308	
Db 156 nPrEaAnPrEaRPhlSPRoSeRPhlARPhaSPARnglu**1leuYsThrIleuYsG1176	
QY 309 CATCGTGAAGGCCAAACCGTGAAGGTCCGACCTATGATTTGTGACACACTCMAAGTT368	
Db 176 uIlETTrgIuNgIYlStmTrAlGlnIllePRoVAlTYARSPheValSeRhlSseARgIy196	
QY 369 ACCAGAGACACAGGTGTCTACCCGCGGAGAGTGGTCTGTGGAGGGCACTTGGGTT428	
Db 196 SgluGluTrAlThrValTYRPRoAlaSPValValleuPheGluIYlIleuAlaPh216	
QY 429 CTACAGCCAGAGATCCGGACATTTTCACTTCGCGCTCTCGTGACACGACTCCGA488	
Db 216 eTYRSeRgInGluValARgARPeuPheGlnMeTLYleuPheValAspThrARPhAlaS236	
QY 489 CGTCAAGCTGTCTGAAAGATTTCTCCGGACGTG--CGCCGAGGAGGACCTGGAGCA545	


```
Db 236 pthArvgLseSerArgrgValleuArGspIleSeSerGluArGlyArGspLeuGlu 256
Qy 546 GATTTCAGCCAGTACACCACTTCGTGAAGCCGCTTCGAGGAGTTTCGCTGCCGAC 605
Db 256 nileleSeSerGlnThrIleThrPheValysProAlaPheGlnGluPheCysLeuProth 276
Qy 606 AAGAAGTATGCGCATGATGATCCAGAGGAGTGAATATGATGATGATGATGATGATGAT 665
Db 276 rlySlyrAlaAlaSpaIleleleProArGlyAlaSpaSmuValAlaIleAsnle 296
Qy 666 GATCGTCAGACATCCAGACATTCGATGATGATGATGATGATGATGATGATGATGATGAT 725
Db 296 uileValGlnHleIleGlnAspIleleuAnGly----- 307
Qy 726 GTCCAAATGGCGGAGCTACCAAGCGGACCTTTCTGAGCCAGGAGCACCTCGGATGCT 785
Db 308 -----GlyProSerIysArGlnThr-----AsnGlyCylleuAnGlyrTrh 322
Qy 786 GACCTCGGCAACGGTCACATTTGAGTCCAGACGAGACCCGAC 831
Db 322 rProSerArGlyArGlnAlaSerGlnSerSerArGProHis 337

RESULT 3
US-09-896-522-5
; Sequence 5, Application US/09896522
; Patent No. US20020055161A1
; GENERAL INFORMATION:
; APPLICANT: Glucksmann, Maria A.
; TITLE OF INVENTION: 57658, A NOVEL HUMAN URIDINE KINASE AND
; FILE REFERENCE: 381552001700
; CURRENT APPLICATION NUMBER: US/09/896,522
; CURRENT FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: 60/216,503
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 125
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Consensus amino acid sequence
US-09-896-522-5

Alignment Scores:
Pred. No.: 5,34e-28 Length: 125
Score: 428.50 Matches: 84
Percent Similarity: 76.80% Conservative: 12
Best Local Similarity: 67.20% Mismatches: 28
Query Match: 28.30% Indels: 1
DB: 9 Gaps: 1

US-09-896-522-3 (1-834) x US-09-896-522-5 (1-125)
Qy 460 CTGGCGCTCTTCGGGACACCGACTCCGAGCTGAGGCTGCTCGAAGATTCCTCCGGAC 519
Db 1 leuLysIlePheValaPthrAspAlaAspValaArGleuIleArGArgIleLysArgAsp 20
Qy 520 GTG---CGCCGAGGAGGAGCTGAGACATTCGAGCAGTACACCACTTCGTAAG 576
Db 21 ValAsnGluArGlyArGAspIleleIleuSerValIleleuGlnIlyrMetClySpheValys 40
Qy 577 CCGGCTTCGAGAGTTCTCGCTCGACAAAGAGTATCCGATGATGATGATGATGATGATGAT 636
Db 41 ProMetIyrGlnGlnPheIleGluProThrIleYrTrhIleAspIleIleIleProArg 60
Qy 637 GAGAGTGAACAATATGCTTCGATCAACCTGATGTCGAGACATCCGACATTCGAAT 696
Db 61 GlyGlyAspAsnHisValAlaIleAspLeuIleValGlnHisIleGlnSerIleLeuAsn 80
Qy 697 GGTCACATTCGAATATGACCGGAGGAGGTCCATGAGGCGGAGCTACAGCGAAGCTTT 756
```

```
Db 81 GluGlyLeuSeSerSerGlnHisThrAsnIyrMetValAsnArGSerIyrIysArGThrPhe 100
Qy 757 TCTGACCGAGGAGGACCACTTCGAGTGCACCTTCGCAACGGTCACATTTGGAGTCC 816
Db 101 SerGlnProGlyAspHisProGlyrTrhProSerGlyLysArGlnHisLeuGlnSer 120
Qy 817 AGCAGCAGACCCGAC 831
Db 121 SerSerArGProHis 125

RESULT 4
US-10-029-386-28765
; Sequence 28765, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR GI
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 28765
; LENGTH: 60
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR9.1
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.5
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.8
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.6
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.7
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.5
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.7
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.3
; OTHER INFORMATION: SWISSPROT HIT: P52623, EVALU8 8.00e-26
US-10-029-386-28765

Alignment Scores:
Pred. No.: 1.86e-19 Length: 60
Score: 327.00 Matches: 60
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 21.60% Indels: 0
DB: 12 Gaps: 0

US-09-896-522-3 (1-834) x US-10-029-386-28765 (1-60)
Qy 652 GTTGCCATCAACCTGATCGTCGACGACATCCAGACATTCGATGATGATGATGATGATGAT 711
Db 1 ValAlaIleAsnleuIleValGlnHisIleGlnAspIleleuAnGlyAspIleCyls 20
Qy 712 TGGCAGCGAGGAGGTCGATGAGGCGGAGCTACCAAGCGGACCTTTCTGAGCCAGGAGGAC 771
Db 21 TrpHisArGlyGlySerAsnGlyArGSerIyrLysArGThrPheSerGlnProGlyAsp 40
Qy 772 CAGCCCGGATGTCGACCTTCGCAACCGTACATTCGATGATGATGATGATGATGATGATGAT 831
Db 41 HisProGlyMetCueThrSerGlyLysArGSerHisLeuGlnSerSerArGProHis 60

RESULT 5
US-09-896-522-6
; Sequence 6, Application US/09896522
; Patent No. US20020055161A1
; GENERAL INFORMATION:
; APPLICANT: Glucksmann, Maria A.
; TITLE OF INVENTION: 57658, A NOVEL HUMAN URIDINE KINASE AND
; FILE OF INVENTION: USES THEREOF
```

```
; FILE REFERENCE: 381552001700
; CURRENT APPLICATION NUMBER: US/09/896, 522
; CURRENT FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: 60/216,503
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 124
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Consensus amino acid sequence
US-09-896-522-6

Alignment Scores:
Pred. No.:      8,17e-19      Length:      124
Score:          320.50        Matches:      62
Percent Similarity: 67.19%    Conservative: 24
Best Local Similarity: 48.44%  Mismatches:    37
Query Match:    21.17%       Indels:       5
DB:             9            Gaps:         2

US-09-896-522-3 (1-834) x US-09-896-522-6 (1-124)

QY 73 CTGATAGGGGTGAGCGGGGACCTGCCGCGGAAAGTCGACCTGTGTGAGAGATCATG 132
DB 1 I l l e l l e g l y l e a l a g l y s e r g l y s t h r l l e a l a g l y s i l e v a l 20
QY 133 GAGTGTCTGGACGACGAAGTGGAAACAGCGGACGCGGACGGAAGTGTGATCTCTAGCCAG 192
DB 21 G l u t e t l e u a n ----- l y e p r o g l y n g l u l y s v a l l a i l l e i l s e r c l n 36
QY 193 GACAGGTTCTTACAAAGTCTCTGACGCGACAGGACGAAAGCCGCTTGAAGAAGACAGTAC 252
DB 37 A s p a n t y r t y l y s a s p l e u s e r c l u d e u a s p m e t g l u g l u a r g l y s l u a s n a n t y x 56
QY 253 AATTTGACCATCCAGATGCTTTGATTAATGATTGAGCAGCAGAGACTCTGAAGACATC 312
DB 57 A s n h e h e a s p h i s p r o a s p a l a p h e a s p h e u l e u t y r g l u n h i s l e u l y s * * L e u 76
QY 313 GTGAGGCGCAAAAGCGTGGAGAGTCCGACCTATGATTTTGGACA---CACTCAAGGTTA 369
DB 77 L y s a n e n g l y s s e r v a l g l u v a l p r o l l e t y r a s p h e u l y s t h r h i s h i a r g d r l y s 96
QY 370 CCAAGACCAACGCGTGTACCTGCGGACGCTGTTCTGTTGAGGCGATCTTGAGTTC 429
DB 97 A s p g l u t h r v a l t h r l l e g l u p r o l a a s p a l l e i l l e u g l u g l y l e t y r a l a l e u 116
QY 430 TACAGCCAGGATCCGGGACATG 453
DB 117 T y r a s p g l u a r g l l e a r g a s p l e u 124

RESULT 6
US-09-896-522-4
; Sequence 4, Application US/09896522
; Patent No. US20020055161A1
; GENERAL INFORMATION:
; APPLICANT: Glucksmann, Maria A.
; TITLE OF INVENTION: 57658, A NOVEL HUMAN URIDINE KINASE AND
; FILE REFERENCE: 381552001700
; CURRENT APPLICATION NUMBER: US/09/896, 522
; CURRENT FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: 60/216,503
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 231
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
```

```
; OTHER INFORMATION: Consensus amino acid sequence
US-09-896-522-4

Alignment Scores:
Pred. No.:      5.95e-15      Length:      231
Score:          276.00        Matches:      68
Percent Similarity: 50.95%    Conservative: 39
Best Local Similarity: 32.38%  Mismatches:    81
Query Match:    18.23%       Indels:       22
DB:             9            Gaps:         5

US-09-896-522-3 (1-834) x US-09-896-522-4 (1-231)

QY 73 CTGATAGGGGTGAGCGGGGACCTGCCGCGGAAAGTCGACCTGTGTGAGAGATCATG 132
DB 1 I l l e g l y l a l a g l y s e r g l y a l a g l y s t h r v a l l a a r g a g l l e v a l 20
QY 133 GAGTGTCTGGACGACGAAGTGG-----GNACAGCGGACG 168
DB 21 S e r l l e p h e d l y a r g l u g l y a l p r o l a a l a g l y l l e g l u d l y a n p r o a s p e r a n 40
QY 169 CGAAGGTGTCTATCTGACCGACGACAGGTTCTTACAGGTTCTTACGCGACGACGAAG 228
DB 41 T h r g l y a s p e r p h e u a r g l e u a s p a r g p h e t y m e t a s p l e u n h i s l e u g l u a s p a r g 60
QY 229 GCCAAGGCTTGAAGAAGACGATCAATTTGACCATCCGATCCGATGCTTGAATAATGATTG 288
DB 61 L y s a r g l a l a l y a n l y s h i s t y r s e r p h e s e r p r o g l u n l a s n a s p h e a s p l e u 80
QY 289 ATCCACAGCACTCTGAAGACATCTGTGAGAGGCGCAACGATGAGGTGCGGACCTATGAT 348
DB 81 L e u t y r g l u v a l p h e y s g l u e u l y s g l u g l y l y s e r v a l a s p y s p r o l l e t y r a n 100
QY 349 TTTGTGACACATCAAGTTTACCAAG-----ACCAAGGTGCTTACCT 393
DB 101 H i e v a l t h r g l y l u a r g a s p r o a s p e r g l y n g l u p r o g l y t h r p h e t h r a s p r t p r o 120
QY 394 -----CGGACGCGGTTCTGTTGAGGCGCATCTGTTGTTCTACACCGAC 438
DB 121 G l u l e u l l e g l u g l y l a s p a l l e u a l l e g l u g l y u e u n i s a l a l e u t y r a s p g l u 140
QY 439 -----GAGATCCGGGACATGTTCCACTCGGCTCTTCTGTGACACCGACTCCGAC 489
DB 141 A r g l u v a l a s n a l a g l u n l e u a s p l e u l y s i l l e t y r a l a s p r o a s p i l l e a s p 160
QY 490 GTGAGCTGTCTGAAGAGTTCTCCGGACGTG---CGCCGAGGAGGACCTTGAGACG 546
DB 161 L e u g l u e u a l a r g y s i l e g l n a r g a s p m e t a l a g l u a r g l y h i s s e r l e u g l u g l y 180
QY 547 ATTCTGACGACGATACACACCTTCGTGAAGCGCGCTCGAGGAGGATTCGCTCCGACGA 606
DB 181 V a l l e u a s p e r l l e g l u l y s a r g a r g l y s p r a s p t y r v a l a n t y r l l e a l a p r o g l n 200
QY 607 AAGAAGTATGCCGATGTGATCATCCACGA 636
DB 201 P h e s e r t y r l a a s p l e u l l e i l l e g l n a r g 210

RESULT 7
US-09-975-719-369
; Sequence 369, Application US/09975719
; Publication No. US20030022349A1
; GENERAL INFORMATION:
; APPLICANT: Ausubel, Frederick M.
; APPLICANT: Rahme, Laurence G.
; TITLE OF INVENTION: VIRULENCE-ASSOCIATED NUCLEIC ACID
; FILE REFERENCE: 00786/361003
; CURRENT APPLICATION NUMBER: US/09/975, 719
; CURRENT FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 09/199,637
; PRIOR FILING DATE: 1998-11-25
; PRIOR APPLICATION NUMBER: US 60/066,517
; PRIOR FILING DATE: 1997-11-25
```

```

; NUMBER OF SEQ ID NOS: 437
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 369
; LENGTH: 265
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-975-719-369

Alignment Scores:
Pred. No.:      4.55e-05      Length:      265
Score:          160.00      Matches:      87
Percent Similarity: 35.62%      Conservative: 22
Best Local Similarity: 28.43%      Mismatches: 107
Query Match:      10.34%      Indels:      90
DB:              11          Gaps:      17

US-09-896-522-3 (1-834) x US-09-975-719-369 (1-265)
QY      820 TGGTGAAGTCCAAATGTGACCGTTTGGCAGAGGTACAGATCCGAGGGTGGTCCCTGGCT 761
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      11  CysArgSerProAlaValProArgCysCysArgSerIleYs--GlyArgAlaAla 29
QY      760 CAGAAAAGTCCGCT-----TGTAGCTCCGCCAT-----731
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      30  GluSerArgSerAlaProTyrProAlaProCysAlaIleAlaGlyAsnProProArgThrCys 49
QY      730 -----TGACCCCTCCGCTGGCATTTGCAGATGTACCATTCAGATGTCTGGATGT 677
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      50  ArgGlySerSerAspLeuSerProValProArgAlaValProGlyAlaAlaSerThr 69
QY      676 GCTGACAGATCAGATGATGAGCAACATATTGTCCATCTCCGAGGATGATCATCGG 617
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      70  ThrArgArgSerArgAlaGlyLeuProProAlaAlaProCys-----Arg 83
QY      616 CATACTTCTTTGTCGCGCAGCAGAACTCTTCGAGGCCGCTTCACGAGGTGGTACT 557
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      84  HisGlySerArgHisAla--ArgSerProArgArgProLeuAspArg--Arg 100
QY      556 GCGTCAGATCTGTCTCCAGAGTCCCTCCGCGCGCAGCTCCCGAGAACTTTCGAGACA 497
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      101  AlaThrSerArgAlaPro--ProSerValHisAla-----111
QY      496 GCGTCAGTGGAGTGGTGTCCACGAAAGGCGCAGTGAACATGTCCCGATCTCT 437
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      112  -----ThrArgAspTyrArgIleValAlaIleGlyeProAlaMetSerAlaPro 127
QY      436 GGGCTGAGAACACCAAGATCCCTCAAAACAGAACACAGTCCGCGAGGTAGACCAACCGTGG 377
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      128  Gly-----GluAspCysValAlaHisTyr--ProArgProThrCysAspArgProPro-- 142
QY      376 TCTCTGGTAACCTTGAAGTGTGTCAAAATCATATGAGTGGCAGCTCCACCGTTTGCCCT 317
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      143  -----AlaGlySerLeuArgProAlaArgAlaProProAlaArgPro 156
QY      316 CCAAGATGTTCTTCAGAGTCTCTGATCAATAATCATATTAAGAAGGATCTGANTGTGCA 257
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      157  ProArgSerAlaLeuProSerThrAla-----GlyCys 167
QY      256 AATTGTAAGTCTCTTCAGAGCGCTTGCGCTTGTGCTCGCG-----TCAGA 209
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      168  GlyCysVal-----ProThrAlaValProArgSerGlyArgProGly 181
QY      208 CCTTGTGAAC-----197
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      182  ProAlaSerThrGlyAlaArgTyrGlyProArgAspArgAlaArgSerArgAlaAsn 201
QY      196  ---TGCTCGCTCAGAGATACCACTTCGCGTGGCTGCTTCACCTGCTGTGTGCCA 140
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      202  AlaAlaProIleProGlyArgSerProAlaThrSerAla--ProGlyArgSerProPro 220
QY      139  GCAACTCCATGATCTTCTCAACACGCGTGCAGCTCCGCGTGGCAGTGGCCGCGCTCACCC 80
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      221  GlyCys-----SerAlaThrAspArgAlaAlaAlaArgAspSerArg 233

```

```

QY      79  CTATCAGAAAGGCGCGCTGTGTGCGACGATCGGCGCTCGGCGCGGCGCTCGAGTCTT 20
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      234  LeuProProArgThrProHisAlaHisSerArgProAlaProProIleSerProAlaVal 253
QY      19  CGCCTCCGCGCAAGCCA 2
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      254  ArgArgAlaProAlaPro 259

RESULT 8
US-09-738-626-4609
; Sequence 4609, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: MAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738, 626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 4609
; LENGTH: 312
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-4609

Alignment Scores:
Pred. No.:      0.00014      Length:      312
Score:          154.50      Matches:      61
Percent Similarity: 44.64%      Conservative: 39
Best Local Similarity: 27.23%      Mismatches: 87
Query Match:      10.20%      Indels:      37
DB:              10          Gaps:      12

US-09-896-522-3 (1-834) x US-09-738-626-4609 (1-312)
QY      67  CCTTCTGATAGGGGTGACCGCGCACTGCGCAGCGGAAGTCAACCGTGTGTGAAG 126
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      90  ProPheValIleGlyValAlaGlySerValAlaValGlySerThrThrAlaGlyLeu 109
QY      127  ATCATGAGATTGCTGGGACAGAAACGAGTGAACACCGG-----CAGCGGAAGGTG 177
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      110  LeuGlnValLeuLeu-----GlnArgTyrAsnSerHisProArgVal 123
QY      178  GTATCCTAGCAGCAGACAGATTTCTACAGAGTCTTACAGGCGACAGCAAGCCAGGCC 237
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      124  AspLeuValThrThrAspGlyPheLeu--TyrProGlyAlaGlyLeuIleArgArgGly 142
QY      238  TTGAAAGACAGTACATTTTGAACATCCAGATGCTTGTATATGATTTGTGACAGG 297
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      143  LeuMetSerArgIleGlySerLyshe-----ProGlySerTyrAspGlnArgAlaLeuLeuArg 160
QY      298  ACTCTGAAGAACATCTCGGAGGCAAA--ACGGTGAGAGTCCCGACCTATGATTTTGTG 354
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      161  PheValThrAspValIleGlySerGlyLeuGlyValAsnAlaProValIleSerHisThr 180
QY      355  ACACACTCAAGGTACCA--GAGACCAAGGTGTACTCCGCGGACGTGTCTGTTT 411

```

```
Db 181 AlaTyrAspArgValProGlyGluPheThrValAlaArgInProAspIleLeuVal 200
Qy 412 GAGGCACTTGGTGTTCACAGCCAGAG-----ATCCGGACATTTCCACCTG 462
Db 201 GluGlyLeuAenValLeuGlnThrGlyProThrLeuMetValSerAspLeuPhe 220
Qy 463 CGCCTCTCGAGCACC---GACTCCGACGTCAGG-----CTGCTCGAAGGTT 510
Db 221 SerValTyrValAspAlaArgThrGluAspIleGluTyrTyrIleAspArgPhe 240
Qy 511 ---CTCCGGACGTGCCCGCAGGAGG----- 534
Db 241 LysLeuArgAspThrAlaPheArgArgProGlyAlaHisPheSerHisTyrAlaAspMet 260
Qy 535 ---GACCTGGAGCAGATTCTGACGAG-----TACACCACTTGTGAACCGGCC 582
Db 261 AlaAspProGluSerIleAlaValAlaArgGluLeuTyrGlnSerIleAsnLeuProAsn 280
Qy 583 TTCGAGAGTCTGCTCCGACAAAGATATGATGATGATCATCCACGAGAGTG 642
Db 281 LeuValGluAenIleLeuProThrArgValArgAlaSerLeuValLeuLysGlySer 300
Qy 643 GACAATATGTT 654
Db 301 AspHisLeuVal 304
```

RESULT 9

```
US-10-156-761-10907
; Sequence 10907, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; PRIOR FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 10907
; LENGTH: 797
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-10907
```

Alignment Scores:

```
Pred. No.: 0.000202 Length: 797
Score: 154.00 Matches: 100
Percent Similarity: 34.12% Conservative: 30
Best Local Similarity: 26.25% Mismatches: 110
Query Match: 9.95% Indels: 141
Gaps: 18
```

US-09-896-522-3 (1-834) x US-10-156-761-10907 (1-797)

```
Qy 811 CCAATGTGACCGTTTGGCAGAGGTACAGATCCAGAGGTGCTCCCTGGCT----- 761
Db 399 ProProAlaSerGlyAlaGlnArgSerValIleCysGlyAlaSerValAlaThrAlaArg 418
Qy 760 -----CAGAAAGGTCCGCTTGATGCTCCGCCCATTTGAGC----- 725
Db 419 ArgThrGlyGlnLysGlyPheSerValSerSerThrArgValThrLeuArgArgAsnGln 438
Qy 724 ---CTCTCGGTGCATTGTCAGATGTCACCATTCAGAAATGCTGATGT----- 677
```

```
Db 439 ArgLeuValTyrPheAlaLeuAlaGlnArgSerHisValProCysAlaGlyCysProGluSer 458
Qy 676 ----- 668
Db 459 ArgLeuArgAlaLeuArgProLeuProSerProSerLeuLeuProAlaAlaThrArg 478
Qy 667 TCAGGTGATGAGGACCATATTGT-----CCATCCCTC 635
Db 479 ProAlaSerProArgProAlaAlaAlaAlaAlaArgProAlaAlaLeuProSerPro 498
Qy 634 GTGGATGATCATCGCATCTGCTTTGTG----- 602
Db 499 ThrThrSerAlaAlaArgAlaThrSerProSerMetThrLeuProProValTyrArg 518
Qy 601 -----CGAGCAGAACTCTCGAAGCCGGCTTCAGAAAG----- 566
Db 519 ArgProAlaArgSerSerArgProProArgProSerSerArgThrAlaSerPro 538
Qy 565 -----TGG-----TGTACTGGCTCAGAA 548
Db 539 ThrProThrArgCysSerValTyrSerSerSerProSerProAlaThrThrArgSerVal 558
Qy 547 TCTGCT---CCAGTCCCTCCCTCGCGCAGCTCCCGAAGACTTTCAGACAGCCTGA 491
Db 559 SerAlaSerPro-ThrProProSerSerArgLysProProArgSerThrArgArgPro-P 578
Qy 490 CGTCGAGTGGGTGTCCAGCAAGAGCGCAGGTGGAACATGTCCCGATTCCTGGCTGT 431
Db 578 roSerAlaSerSerThrThrSerArgSerArgThrThrSerArgThrSer-SerSer 597
Qy 430 AGAACACCAAGATGCCCTCAACAGAACACGCTCCGAGGTGACACCG----- 380
Db 598 ThrArgArgArgProThrThrSerProAlaSerProProArgProProSerArgThr 617
Qy 379 -----TGTCTCTGTGAACCTTGA----- 359
Db 618 SerSerAlaSerSerAlaAlaTyrThrSerArgSerThrSerSerArgAlaSerArg 637
Qy 358 -----GTGTCAAAATCATAGTGGCGGACCTCCACCGTTT 323
Db 638 ArgAlaSerArgThrProArgArgValSerArgSerSerArgSerThrProArgPro-- 656
Qy 322 TGGCCTCCAGCATGTTCTTCAGAGTCTCGTCATCAATCATATCAAGAGCATGTGAT 263
Db 657 ---ProProArgAla----- 600
Qy 262 GGTCAAAATTTGATCTGCTTTCAGAGCCTTGACCTTCTGCT----- 221
Db 661 GlySerProAlaArgThrArgAlaArgProArgProArgAlaArgSerAlaProArgPro 680
Qy 220 CTGCGCTCAGAGACTGTAGAACTGCTCTGCTCAGATGACCACTTCGCTGCCGCT 161
Db 681 ThrSerSerThrProProArgAlaCysProAla-----ArgAlaSerSer 695
Qy 160 GTTCCACT-----CGTTCTGCCCACTCATCATGATCTTCTCAC 119
Db 696 ArgProProArgProArgThrArgCysArgArgSerVal---SerThrProThrSerThr 713
Qy 118 ACAAGGTGACCTTCCGCTGCGAGTCCGCGCTCAACCCCTATCAGAAAGCGCGCTGCT 59
Db 714 -----SerArgThrArgTyrArgSerThrArgThrProSerSerProArgArgArgThr 731
Qy 58 GCGAGAGGTGCG-----CCTCGGCGCGGCGGCTTCGACGTTCGCTCCCGC 11
Db 732 TrrProAlaArgCysThrThrSerProSerArgCysThrThrAlaSerArgArgPro 750
```

RESULT 10

```
US-10-156-761-12495
; Sequence 12495, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
```

```

; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 12495
; LENGTH: 329
; TYPE: PRT
; ORGANISM: Streptomyces avermectilis
US-10-156-761-12495

Alignment Scores:
Pred. No.: 0.000417 Length: 329
Score: 149.00 Matches: 62
Percent Similarity: 43.67% Conservative: 38
Best Local Similarity: 27.07% Mismatches: 89
Query Match: 9.84% Indels: 40
DB: 15 Gaps: 11

US-09-896-522-3 (1-834) x US-10-156-761-12495 (1-329)
QY 67 CCCTTCCTGATAGGGGTGAGGGCGGCGCATGCGCGGGAAGTGCAGCTGTGTGAGAG 126
Db 100 ProPheValIleGlyValAlaGlySerValAlaValGlySerThrValAlaArgLeu 119
QY 127 ATCATGAGTGTGCTGGAGCAAGAGGATGGAACAGCGGCGGAGCGGAGTGTGATCCTG 186
Db 120 LeuGlnAlaLeuLeuSerArgTrpProGlnHisProArg-----ValGlnLeuVal 136
QY 187 AGCCAGACAGAGGTTCTACAGAGTCTTACGCGCAGACAGCAAGAGCCCTGAAAGGA 246
Db 137 ThrThrAspGlyPheLeu---LeuProThrIleGlyLeuGlnAlaArgGlyLeuMetSer 155
QY 247 CAGTACCAATTTGACCATCCAGATGCTTGATATATGATTGACGACAGACTCTGTAAG 306
Db 156 ArgIleGlyPhe-----ProGlnSerTrpAspArgArgAlaLeuThrArgPheValAla 173
QY 307 AACATCGTAGAGGGCAAA---ACGCTGAGAGGCGGACCTATGATTTTGTGACACACTCA 363
Db 174 AspIleValAlaGlyIleAspGlnValThrAlaProValIleTrpSerHisLeuIleTrpAsp 193
QY 364 AGGTATCA---GAGACCAAGGAGTGTACCTTCGCGGACGTGTTCTGTTGAGGGC 417
Db 194 IleValProGlnGlnIleuThrValArgArgPro---AspIleLeuIleValGlnGly 212
QY 418 ATCTTGTTGTTCTACACCGA---
Db 213 LeuAsnAlaLeuGlnProAlaLeuProGlyIleAspIleValArgThrArgValAlaGlyLeuAla 232
QY 448 GACATGTTCCACCTGCGCCTCTTGTGTAACAC---GACTCCGAGTGCAGG----- 495
Db 233 AspTrpThrPheAspPheSerValIleValAlaAspAlaArgProGlnAspIleGlnIleGlyTrpTyr 252
QY 496 CTGTCTCGA---AGAGTTCTCCGG----- 516
Db 253 LeuAsnArgPheArgIleValArgAlaThrAlaPheGlnAsnProSerSerTrpPheArg 272
QY 517 -----GACGTGCGCCGAGGAGGAGGAGCTGAGACAGATTCGAGCGAGTACACACC 567
Db 273 ArgGlyThrGlnValSerGlnAspGlnAlaLeuAspTrpAlaArgGlnThrTrpArgThr 292
QY 568 TTGCGAAGCGGCGCTTCGAGAGTTCGCTCGCGCAAAAGATATGCGGATGTATC 627

```

```

Db 293 IleAsnIleValAsnLeuLeuGlnAsnValAlaProThrArgGlyArgAlaIleLeuVal 312
QY 628 ATCCACAGAGAGTGTGACATATAGTT 654
Db 313 ValArgIleGlyProAspHisIleVal 321

RESULT 11
US-10-078-547-2
; Sequence 2, Application US/10078547
; Publication No. US2002019211A1
; GENERAL INFORMATION:
; APPLICANT: Narayanaswamy Ramesh
; APPLICANT: Miguel A. de la Fuente
; APPLICANT: Ines M. Anton
; APPLICANT: Raif S. Geha
; TITLE OF INVENTION: WIP, A WASP-Associated Protein
; FILE REFERENCE: 1242.1022-005
; CURRENT APPLICATION NUMBER: US/10/078,547
; CURRENT FILING DATE: 2002-02-19
; PRIOR APPLICATION NUMBER: 09/599,287
; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: PCT/US98/27501
; PRIOR FILING DATE: 1998-12-22
; PRIOR APPLICATION NUMBER: 60/101,457
; PRIOR FILING DATE: 1998-09-23
; PRIOR APPLICATION NUMBER: 60/068,533
; PRIOR FILING DATE: 1997-12-23
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 503
; TYPE: PRT
; ORGANISM: Human
US-10-078-547-2

Alignment Scores:
Pred. No.: 0.00303 Length: 503
Score: 139.50 Matches: 83
Percent Similarity: 33.10% Conservative: 10
Best Local Similarity: 29.54% Mismatches: 133
Query Match: 9.02% Indels: 55
DB: 14 Gaps: 13

US-09-896-522-3 (1-834) x US-10-078-547-2 (1-503)
QY 801 CCGTTTGCAGAGGTGACATCCAGGGTGTGCTGCTCGAAGAGGTCCGCTTGTGA 742
Db 193 ProIleGlnSerSerLeuHisAsnArgGlySerProProValProGlyIleProArgGln 212
QY 741 GCTCCGCCCATTTGACACCTTCCTCGGTGCCATTTTGCAGATGTCAACATTCAGATTCCTG 682
Db 213 ProSerPro---GlyProThrProProProPhe----- 222
QY 681 GATGTGTGCACATGAGTTGATGGAACCATATTGTCCACTCTCGTGGATGATCAC 622
Db 223 -----ProGlyAsnArg---GlyThrAlaLeuGlyIleVal 235
QY 621 ATGCGCATCTTCTTTGTG-----GACAGCAGAACTCTCGAAGCGCG 577
Db 235 IeArgIleSerProLeuSerSerSerProPheSerHisArgProProLeuProProT 255
QY 576 CTTCAGAGAGGTGTGTATGCTGCTCGATCTGCTCCAGGTCCCTCGGCGGACGTC 517
Db 255 hrProSerArgAlaLeuAspArgIleProProProProProProValGlyAsnArg 275
QY 516 CCGGAGAACTTCGAGACAGCTGACGTCGAGTGGGTGTCGACGAGAGGCGGAGGTG 457
Db 275 ro-----SetIleHisArgGlnAlaValProProProProProGln 289
QY 456 GAAATGTCCTCGGATCTCTGCTGTGAGAACACCAAGATGCTCAAGAACACGATC 397
Db 289 snAsnIleProProValProSer-----ThrProArg---ProSerAlaProHisArg 306

```


US-10-078-547-24

Alignment Scores:

Pred. No.: 0.0132 Length: 507
 Score: 132.00 Matches: 79
 Percent Similarity: 31.29% Conservative: 13
 Best Local Similarity: 26.87% Mismatches: 124
 Query Match: 8.53% Indels: 78
 DB: 14 Gaps: 12

US-09-896-522-3 (1-834) x US-10-078-547-24 (1-507)

```

QY 801 CCGTTTCCAGAGTGCAGATCCAGAGGTGATCCCTGCTCAGAAAGGCGCTTGT 742
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 196 ProlleInserleuHisAsnArgGlySerProProValProGlyGlyProArgGln 215
QY 741 GCTCCGCCATTTGAGACCTCTCGGTGCCATTGTCAGATGTCACCATTCAGAAATGTCCTG 682
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 216 ProSerPro---GlyProThrProProPhe----- 225
QY 681 GATGTGCTGCACATCAGTTGATGCAACCATATTTCCTCTGCGGATGATCAC 622
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 226 -----ProGlyAsnArg---GlyThrAlaLeuGlyGlySerI 238
QY 621 ATCGGCATATCTTTGTCG-----GCAGGAGAACTCCTCGAAGGCGCG 577
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 238 leArgInserProleuSerSerSerSerProPheSerAsnArgProProleuProProT 258
QY 576 CTTACGAGAGTGTGATGCGGTGAGATCTGCTCCAGGTCCCTCCGCGCAGCTC 517
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 258 hrProSerArgAlaLeuAspAspIysProProProProProProValGlyAsnArgP 278
QY 516 CCGGAGAACTCTTCGAGACGCTGAGTGGAGTCCGCTGCCAGAAAGGCGCAGTGC 457
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 278 ro-----SerIleHisArgGluAlaValProProProProGlnA 292
QY 456 GAACATGTCCCGGATCTCTGGCTGTGAAACACCAAGATCCCTCAAAACAGAACACGCTC 397
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 292 snAsnIysProProValProSer-----ThrProArg---ProSerAlaProHisArgP 309
QY 396 CGCAGAGTAGAACCCGCTGTCTCTGTGTAACCTTGAGTGTGTCAAAATCATAGTACGCG 337
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 309 roHisLeuArgProPro-----ProProSerArgProG 320
QY 336 CACCTCCACCGCTTTTGCTTCCAGATGTTCTTCAGAGTCTGTGCATCAAAATCATATATC 277
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 320 IyProProProleu---ProProSerSerSerGlyAsnAspGluThrProArgLeuProG 339
QY 276 AAAGCATCTGATGTGCAAAATGTACT-----G 247
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 339 InArgAsnLeuSerLeuSerSerSerThrProProleuProSerProGlyArgSerGlyP 359
QY 246 TCCCTTCAGAGCCTTGCGCTTGTGCTGCGGTGAGAGCCTTGAGAACCTGTGCTGGCT 187
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 359 roLeuProProProValProSerGlnArgProProProValArgAspProProGlyA 379
QY 186 CAGAGTACCACTTCGCGTCCGCTGTTCACCT----- 152
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 379 rgSerGlyPro-----LeuProProProProProValSerArgAsnIys 394
QY 151 -----CGTTGTGTCAGAACCTCATGATCTTCTACACACGCGTGCATTCGCGCT 100
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 394 erThrSerArgAlaLeuProAlaThrPro-----GlnL 405
QY 99 GGCAGTGCAGCGCTGCACCCCTATGAGAAAGGCGGTGTGCGAGCGTGGCGCTTCGG 40
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 405 euProSerArgSerGlyValAspSerProArgSerGlyProArgProProleuProProA 425
QY 39 CGGCGGCTCTCGAGTCTTCGCTCCGCGCGGCGAAGCA 2
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 425 spArgProSerAlaGlyAlaProProProProPro 437

```

RESULT 14

US-10-081-872-208

Sequence 208, Application US/10081872

```

; Publication No. US20030125534A1
; GENERAL INFORMATION:
; APPLICANT: Callen, Walter
; APPLICANT: Richardson, Toby
; APPLICANT: Frey, Gerhard
; APPLICANT: Short, Jay M.
; APPLICANT: Mathur, Eric J.
; APPLICANT: Gray, Kevin A.
; APPLICANT: Kerveno, Janne S.
; APPLICANT: Slupska, Malgorzata
; TITLE OF INVENTION: ENZYMES HAVING ALPHA AMYLASE ACTIVITY
; TITLE OF INVENTION: AND METHODS OF USE THEREOF
; FILE REFERENCE: 09010-108001
; CURRENT APPLICATION NUMBER: US/10/081,872
; CURRENT FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/270,495
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: US 60/270,496
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: US 60/291,122
; PRIOR FILING DATE: 2001-05-14
; NUMBER OF SEQ ID NOS: 321
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 208
; LENGTH: 439
; TYPE: PRT
; ORGANISM: Environmental
; US-10-081-872-208

```

Alignment Scores:

Pred. No.: 0.017 Length: 439
 Score: 130.50 Matches: 69
 Percent Similarity: 35.38% Conservative: 23
 Best Local Similarity: 26.54% Mismatches: 99
 Query Match: 8.44% Indels: 69
 DB: 15 Gaps: 16

US-09-896-522-3 (1-834) x US-10-081-872-208 (1-439)

```

QY 709 TGCAGATGTACCATTCAGAAATGCTCTGATGTGTGACAGATCAGATTGANTGCAACCA 650
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 182 CysSerCysGlnArgProArgProSerCys---SerArgAlaGlySerArgAspPro 200
QY 649 TATGTCACTCTCTCGTGGATGATCATCGGATATCTTTGTGCGCAGGAGCAACT 590
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 201 Trp-----ArgArgSerSerArgProSerGly----- 209
QY 589 CCTCGAAGCCGCGCTTCAGAAAGTGTGTAATCGCTCAGATTCGCTCA----- 539
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 210 -----ValHisGlnArgTrpCysProSerThrArgGlnArgProSerArgPro 225
QY 538 -----GGTCCC-----TCCCTCGGCGACGCTCC 515
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 226 ThrSerAlaSerProArgProThrLeuArgGlyProSerArgSerGlnSerAlaArgHis 245
QY 514 GAGAACTTTTGAGACAGCTGACGTGAGT-----CGGTGTCAGAAAGGCGCA 461
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 246 GlnArgArgCysSerLeuGlyArgArgArgSerSerHisArgSerProArgAlaSerAla 265
QY 460 GGT-----GGAACATGTCCCGGATCTCCGTGCTGTGTAACCAACATGC 416
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 266 GlyProSerSerSerArgGlyLeuCysLeuGlySer-----LeuGlnMetCys 281
QY 415 CCTCAA--ACAGAACACAGTCCGACAGGT--AGACACCGGTGTCTGTGTAACCTTG 362
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 282 ProArgHisSerThrProArgTrpArgGlySerArgGlySerTrpGlnTrIleCys--- 300
QY 361 AGTGTCTCAAAATCATAGTGTGCGACCTTCACCGCTTTTGCCCTTCAGATGTTTCTTA 302
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 301 -----ProArgProProleuArgSerProSerArgCysSerPro 313

```

```

QY 301 GAGTCCTGTCATCAATCATATTATCAAGCATCTGATGTCATAAATTGACTGCTT 242
Db 314 GlnArgThrIleGlySerThrArgGlyLeuArgLeuArgGlyLeuArgCys----- 330
QY 241 TCAGAGCCTTGACCTTCTGCTCTGCGTCAGACCTTTAGAACCTGTCTGCGTCAGGA 182
Db 331 -----ProLeuProLeuCysArgArgHisGlyProCysLeuSerCysSerArgAlaPro 348
QY 181 TGACCACTTCGCGCGCTGTTCCACCTGCTGTCTGTCACCAAGAACCTCATGATCTTCT 122
Db 349 AlaTrpSerGlnSerAlaSerLeuPro-----PhePro-----SerGly 361
QY 121 CACACACGCTGACTTCCGCTGCGATGTCGCGCTCACCCCTATCAGAGAGCGCGCT 62
Db 362 ArgThrHisArgGlyGlnArgSerArgArgGlyArgSerProSerArgArgAlaPro 381
QY 61 GGTGCGACGCTGCGCTCCGCGCGGGGCTCTGCGACTTTCGCTC--CCGCCGAG 5
Db 382 -----CysProCysSerProGlyGlnSerIleTrpArgIlePheProArg 397

```

RESULT 15

```

US-10-024-368-5
; Sequence 5, Application US/10024368
; Publication No. US20030027300A1
; GENERAL INFORMATION:
; APPLICANT: THOMPSON, Catherine C.
; TITLE OF INVENTION: HUMAN HAIRLESS GENE AND PROTEIN
; FILE REFERENCE: Thompson-20263/0243435
; CURRENT APPLICATION NUMBER: US/10/024,368
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/287,354
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-07
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/080,888
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 1207
; TYPE: PRT
; ORGANISM: Rat
US-10-024-368-5

```

```

Alignment Scores:
Pred. No.: 0.0305 Length: 1207
Score: 129.00 Matches: 67
Percent Similarity: 33.33% Conservative: 14
Best Local Similarity: 27.57% Mismatches: 72
Query Match: 8.34% Indels: 90
DB: 15 Gaps: 14

```

US-09-896-522-3 (1-834) x US-10-024-368-5 (1-1207)

```

QY 583 AGGCCGGCTTACGAAGGTGTACTGCGTCAGATCT----- 545
Db 241 ArgProAla---LysGluProLeuAlaAlaSerGlnSerGlyMetLeuGlyLeuAlaPro 259
QY 544 -----GCTCAGATGCTCCCTCGGCGACGTCGCGGA 512
Db 260 GlyGlyHisLeuGlnGlnAlaCysAspAlaGluGlyProSerLeuHisGlnArgAspGly 279
QY 511 GAACCTTCGAGACAGCTGAGTCGAGTGGTGTCCACGAGAGCGCAGGT----- 458
Db 280 Glu-----ThrGlyAlaGlyArgGln 286
QY 457 ---GGAACATGTCCCGATCTCTGCTGTAGAACACCAAGATGCTCTCAACAGAACCA 401
Db 287 GlnAsnLeuCysProValPheLeuGly-----TyrProAspThrValPro 301
QY 400 CGTCCG-----CAGGATAGACACCGTGTCTCTGGAACCTTG----- 362
Db 302 ArgThrProTrpProSerCysProProGlyLeuValHisThrLeuGlyAsnValTrpAla 321
QY 361 -----AGTGTGTCAAAATCATATAGTCGGACCTCCA-----CCGTTTTCGCT 317

```

```

Db 322 GlyProGlySerAsnSerPheGlyTyrGlnLeuGlyProProValThrProArgCysPro 341
QY 316 -----CCAGAGATTTCTTCAAGATCTCTGTCATCAATCATATTCAAGACATCTG 266
Db 342 SerProGlyProProThrProProGlyGlyCysCysSerSerHis----- 356
QY 265 GATGTCATAAATTGACTGTCTTTCAGAGCCTTGACCTTCTGCTGCGCTCA----- 212
Db 357 -----LeuProAlaArgGln 361
QY 211 -----GACCTGTAGAACCTGTCTGCTCAGATGATGACCACTTCGCTGCGCT 161
Db 362 GlyAspProGlyProCysArgGlyCysGlnAspSer-----ProGlnGlySerSer 379
QY 160 GTTCCACTGCTGTCTGCCA-----GCACTCATGATCT 125
Db 380 GlyProGlyGlnSerSerGlnArgAsnLysAlaGlySerArgAlaSerProSer 399
QY 124 TCTCACAACAGTTCGACTTCCGCTGAGTCGCGCTCACCCCTATCAGAGAGGCC 65
Db 400 HisHisThrIleLeuLysLeuLysIleThrTrpLeuThrArgHisSerGlnInPheGluCysPro 419
QY 64 GCTGTGCGGACCGTCCG-----CCTCCGCGCGGGGCTTTCGACGCTTTCGCTCCG 11
Db 420 GlyGlyCysProGlyLysGlyGlnSerProAlaThrGlyLeuArgAlaLeuLysArgAla 439
QY 10 CCGAAGCCA 2
Db 440 GlySerPro 442

```

Search completed: November 25, 2003, 07:50:36
Job time : 98.057 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: November 25, 2003, 07:27:32 ; Search time 10.688 Seconds
(without alignments)
6603.180 Million cell updates/sec

Title: US-09-896-522-3

Perfect score: 1514

Sequence: 1 atggctctgcggcgagcgca.....ccagcagcagaccacctga 834

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 657434

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US09896522/rnat_21112003.184106.2885/app_query.fasta.1.2830
-DB=Issued Patents AA -QFMT=fastaan -SUFFIX=n2p.ra1 -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09896522.@CEN_1.1.29.@rnat_21112003.184106.2885 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents AA.*
1: /cgn2_6/prodata/2/iaa/5A.COMB.pep.*
2: /cgn2_6/prodata/2/iaa/5B.COMB.pep.*
3: /cgn2_6/prodata/2/iaa/6A.COMB.pep.*
4: /cgn2_6/prodata/2/iaa/6B.COMB.pep.*
5: /cgn2_6/prodata/2/iaa/PCTUS.COMB.pep.*
6: /cgn2_6/prodata/2/iaa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysts of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1363	90.0	260	4	US-09-536-647-2	Sequence 2, Appl1
2	1325	87.5	277	4	US-09-536-647-3	Sequence 3, Appl1
3	451	29.8	216	4	US-09-134-001C-3618	Sequence 3618, Ap
4	434	28.7	210	4	US-09-107-532A-5024	Sequence 5024, Ap
5	412	27.2	222	4	US-09-198-452A-783	Sequence 783, App
6	164	10.6	1150	4	US-09-252-991A-24671	Sequence 24671, A
7	160	10.3	325	4	US-09-199-637A-369	Sequence 369, App
8	158.5	10.2	274	4	US-09-252-991A-25569	Sequence 25569, A
9	157.5	10.2	247	4	US-09-252-991A-26899	Sequence 26899, A
10	157	10.4	664	4	US-09-252-991A-31116	Sequence 31116, A
11	156	10.1	335	4	US-09-252-991A-24899	Sequence 24899, A
12	156	10.1	395	4	US-09-252-991A-22433	Sequence 22433, A

c 13	155.5	10.1	265	4	US-09-252-991A-21436	Sequence 21436, A
c 14	153	9.9	663	4	US-09-252-991A-30843	Sequence 30843, A
c 15	152.5	10.1	582	4	US-09-252-991A-24055	Sequence 24055, A
c 16	151.5	9.8	472	4	US-09-252-991A-31978	Sequence 31978, A
c 17	150.5	9.7	272	4	US-09-252-991A-31371	Sequence 31371, A
c 18	150	9.7	266	4	US-09-252-991A-19128	Sequence 19128, A
c 19	149.5	9.7	258	4	US-09-252-991A-22452	Sequence 22452, A
c 20	149.5	9.7	511	4	US-09-252-991A-26078	Sequence 26078, A
c 21	148.5	9.6	190	4	US-09-252-991A-25320	Sequence 25320, A
c 22	148.5	9.6	300	4	US-09-252-991A-27679	Sequence 27679, A
c 23	148.5	9.6	467	4	US-09-252-991A-18296	Sequence 18296, A
c 24	148	9.6	335	4	US-09-252-991A-24046	Sequence 24046, A
c 25	148	9.8	375	4	US-09-252-991A-31128	Sequence 31128, A
c 26	148	9.6	710	4	US-09-252-991A-24946	Sequence 24946, A
c 27	147.5	9.5	341	4	US-09-252-991A-23424	Sequence 23424, A
c 28	147.5	9.5	568	4	US-09-252-991A-23264	Sequence 23264, A
c 29	146	9.4	316	4	US-09-252-991A-25345	Sequence 25345, A
c 30	146	9.6	335	4	US-09-252-991A-23948	Sequence 23948, A
c 31	145	9.4	187	4	US-09-252-991A-19743	Sequence 19743, A
c 32	144.5	9.5	316	4	US-09-252-991A-22720	Sequence 22720, A
c 33	144.5	9.5	763	4	US-09-252-991A-30146	Sequence 30146, A
c 34	143	9.2	200	4	US-09-252-991A-21290	Sequence 21290, A
c 35	143	9.4	441	4	US-09-252-991A-24396	Sequence 24396, A
c 36	142.5	9.2	335	4	US-09-252-991A-23674	Sequence 23674, A
c 37	142	9.2	228	4	US-09-252-991A-28381	Sequence 28381, A
c 38	142	9.2	376	4	US-09-252-991A-23577	Sequence 23577, A
c 39	142	9.2	398	4	US-09-252-991A-26217	Sequence 26217, A
c 40	142	9.4	441	4	US-09-252-991A-28965	Sequence 28965, A
c 41	141.5	9.1	228	4	US-09-252-991A-28783	Sequence 28783, A
c 42	141.5	9.1	239	4	US-09-252-991A-25387	Sequence 25387, A
c 43	141.5	9.1	325	4	US-09-252-991A-26580	Sequence 26580, A
c 44	141.5	9.1	369	4	US-09-252-991A-25394	Sequence 25394, A
c 45	141.5	9.1	433	4	US-09-252-991A-28695	Sequence 28695, A

ALIGNMENTS

RESULT 1
US-09-536-647-2
; Sequence 2, Application US/09536647
; Patent No. 6579708
; GENERAL INFORMATION:
; APPLICANT: Johnson, Randall
; TITLE OF INVENTION: No. 6579708e1 Human Uridine Kinase
; FILE REFERENCE: GP50020
; CURRENT APPLICATION NUMBER: US/09/536,647
; CURRENT FILING DATE: 2000-03-27
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 260
; TYPE: PRT
; ORGANSIM: Human
US-09-536-647-2

Alignment Scores:
Pred. No.: 1.64e-130
Score: 1363.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 90.03%
DB: 4
Gaps: 0

US-09-896-522-3 (1-834) x US-09-536-647-2 (1-260)

QY 1 ATGGCTTGGCGGAGCGAAGCTGCGAGCGCCGCGAGCGCGACCGTCGCGAC 60
DB 1 MetAlaSerAlaGlyValGlnAspCysGlnSerProAlaProGluAlaAspArgProHis 20
QY CAGCGGCGCTTCCTGATAGAGGTGAGCGGCGGCGACCTGCGAGCGGAGTGCACCGTGTGT 120
DB 21 GlnArgProPheLeuIleGlyValSerGlyThrAlaSerGlyLysSerThrValCys 40

```
QY      121 GAGAAATCATGAGTTGCTGGGACAGAACGAGTGGACAGCCGCGAGAGGTGTC 180
Db      41 GlnlysllewtcgluleuLeuGlyGlnasnlvvalGluGlnArgGlnArgGlySvalVal 60
QY      181 ATCTGAGCCGAGCAGGTTCTACAAGGTCTCTGACGGCAGAGAGGACGAAAGCCCTTG 240
Db      61 lIleuSerGlnmaphrpheryrlvsvalIleuthralaglGlnGlnSvalalysalaleu 80
QY      241 AAAGACAGTACAAATTTTGACCATCCAGATGCTCTTGATATGATTTGATGACAGACT 300
Db      81 llyeGlyGlnTyraenPheasrphIsProasrphIaPheasrphasrphleuMetHIsArgThr 100
QY      301 CTGAAAGAACCTCGTGGAGGGCAAAACGGTGGAGGTGCCGACCTTATGATTTTGTGACAC 360
Db      101 leuylsasnllevalGluGlyLysThrValGluValProThrTyraasrphValThrHis 120
QY      361 TCAAGTTTACAGACGACGAGTGTACCTCGGACATGTTCCACTGCGCTCTTGTGACAC 420
Db      121 SerArgLeuProGlnThrThrValValTyrrProalAsrphValIleuPheGluGlyIle 140
QY      421 TTGGTGTCTTACAGCCAGAGATCCGGACATGTTCCACTGCGCTCTTGTGACAC 480
Db      141 leuValPheTyrrSerGlnGlnIleArgAsrphMetPheHIsleuArgLeuPheValAspThr 160
QY      481 GACTCCGACGTCAGGCTGTCTCGAAGAGTTCTCGGACGTCGCGGACGCGGAGGAGGACCTG 540
Db      161 AsrSerAspValaArgLeuSerArgValaIleuArgAspValaArgaGlyValaAspLeu 180
QY      541 GAGCAGATTCTGACGACGATACACACTTCGTGAAGCGGCTTCGAGAGTTGTGCTG 600
Db      181 GlnGlnIleleuthrGlnTyrrThrThrPheValIysProalAphGlnGlnPheCysLeu 200
QY      601 CCGACAAAGAAGTATGCCGATGTGATCATCCACGAGAGTGGACAAATATGTTGCCATC 660
Db      201 ProThrLysTyrrAlaAspValIleIleProArgGlyValaAspAsnMetValaIle 220
QY      661 AACCTGATTCGTCGACGACATCCAGACATTTGTAATGTGACATCTGCAATATGACGCA 720
Db      221 AsnleuIlevalGlnHIsIleGlnAsrphIleuasnGlyAsrphIleCysLysrTPHIsArg 240
QY      721 GAGAGGTCCAATGGGCGGAGCTACAGCGGACCTTTCTGAGCGAGGAGGACCACTGAG 780
Db      241 GlynlyserasnGlyArgserTyrrLysArgThrPheSerGlnProGlyAsrphIsProGly 260

RESULT 2
US-09-536-647-3
; Sequence 3, Application US/09536647
; Patent No. 6579708
; GENERAL INFORMATION:
; APPLICANT: Ho, Yen Sen
; APPLICANT: Johnson, Randall
; TITLE OF INVENTION: No. 6579708el Human Uridine Kinase
; FILE REFERENCE: GP50020
; CURRENT APPLICATION NUMBER: US/09/536,647
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 277
; TYPE: PRT
; ORGANISM: Human
US-09-536-647-3

Alignment Scores:
Pred. No.: 1,27e-126 Length: 277
Score: 1325.00 Matches: 254
Percent Similarity: 95.67% Conservative: 11
Best Local Similarity: 91.70% Mismatches: 12
Query Match: 87.52% Indels: 0
Gaps: 0
US-09-896-522-3 (1-834) x US-09-536-647-3 (1-277)
```

```
QY      1 ATGGCTTCCGCGGAGGACCGAAGACTCGAGAGCCCGCGCGGACCGTCCGCAC 60
Db      1 MetAlaSerAlaGlyGlyGlySerGlnSerAlaIleProGlnAlaAsrphProGln 20
QY      61 CACGCGCCCTCTCTGATAGAGGAGTGAACGGCGGACCTCGACGCGGAAAGTGCAGCTGTGT 120
Db      21 ProArgProPheleuIleGlyValSerGlyGlyThrAlaSerGlyLysSerThrValCys 40
QY      121 GAGAAATCATGAGTTGCTGGGACAGAACGAGTGGACAGCCGCGAGAGGTGTC 180
Db      41 GlnlysllewtcgluleuLeuGlyGlnasnlvvalAsrArgArgGlnArgLysleuVal 60
QY      181 ATCTGAGCCGAGCAGGTTCTACAAGGTCTCTGACGGCAGAGAGGACGAAAGCCCTTG 240
Db      61 lIleuSerGlnmaphrpheryrlvsvalIleuthralaglGlnGlnSvalalysalaleu 80
QY      241 AAAGACAGTACAAATTTTGACCATCCAGATGCTCTTGATATGATTTGATGACAGACT 300
Db      81 llyeGlyGlnTyraenPheasrphIsProasrphIaPheasrphasrphleuMetHIsArgThr 100
QY      301 CTGAAAGAACCTCGTGGAGGGCAAAACGGTGGAGGTGCCGACCTTATGATTTTGTGACAC 360
Db      101 leuylsasnllevalGluGlyLysThrValGluValProThrTyraasrphValThrHis 120
QY      421 TTGGTGTCTTACAGCCAGAGATCCGGACATGTTCCACTGCGCTCTTGTGACAC 480
Db      121 SerArgLeuProGlnThrThrValValTyrrProalAsrphValIleuPheGluGlyIle 140
QY      481 GACTCCGACGTCAGGCTGTCTCGAAGAGTTCTCGGACGTCGCGGACGCGGAGGAGGACCTG 540
Db      141 leuValPheTyrrSerGlnGlnIleArgAsrphMetPheHIsleuArgLeuPheValAspThr 160
QY      161 AsrSerAspValaArgLeuSerArgValaIleuArgAspValaGlnArgGlyValaAspLeu 180
QY      181 GlnGlnIleleuthrGlnTyrrThrThrPheValIysProalAphGlnGlnPheCysLeu 200
QY      201 CCGACAAAGAAGTATGCCGATGTGATCATCCACGAGAGTGGACAAATATGTTGCCATC 260
Db      201 ProThrLysTyrrAlaAspValIleIleProArgGlyValaAspAsnMetValaIle 220
QY      221 AACCTGATTCGTCGACGACATCCAGACATTTGTAATGTGACATCTGCAATATGACGCA 280
Db      221 AsnleuIlevalGlnHIsIleGlnAsrphIleuasnGlyAsrphIleCysLysrTPHIsArg 240
QY      241 GAGAGGTCCAATGGGCGGAGCTACAGCGGACCTTTCTGAGCGAGGAGGACCACTGAG 300
Db      241 GlynlyserasnGlyArgserTyrrLysArgThrPheSerGlnProGlyAsrphIsProGly 320
QY      261 ValIleuAlaThrGlyLysArgserHIsleuGlnLysSerSerArgProHIs 277

RESULT 3
US-09-134-001C-3618
; Sequence 3618, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
```

/ SEQ ID NO 3618
 / LENGTH: 216
 / TYPE: PRT
 / ORGANISM: Staphylococcus epidermidis
 / US-09-134-001C-3618

Alignment Scores:

Pred. No.:	1,49e-37	Length:	216
Score:	451.00	Matches:	91
Percent Similarity:	63.68%	Conservative:	44
Best Local Similarity:	42.92%	Mismatches:	61
Query Match:	29.79%	Indels:	16
	4	Gaps:	4

US-09-896-522-3 (1-834) x US-09-134-001C-3618 (1-216)

```

QY 73 CTGATNAGGGTACGCGGCGGCACTGCCAGGGAAGTCACCGTGTGAGAAATCATG 132
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 15 lilelglylealaglyserglyserglyserthrvalthrsealalemet 34
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 133 GAGTTGCTGGAGACGAACGAGTGAACAGCGGACGGGAGAGTGTGATCTTGAGCCAG 192
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 35 Lysasmeu-----Gluglyhiseryalaleuleuallagln 47
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 193 GACAGGTTCTACAAG-----GTCTGACGCGACAGGACGAGGACGCTTGAA 243
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 48 AsptlyrtyrlylsasprglnserhisleuthrPheglugluargleuyluthr----- 65
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 244 GACAGTACAATTTTGACATCCAGATGCTTTGATATGATTTGATGACACAGACTTG 303
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 66 -----AsnlyrasphisprophealaPheasprasnleuilehisasmeu 82
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 304 AATAACATCGTGAAGGCAAAACGCTGAGAGTGCACCACTTATGATTTTGACACACTCA 363
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 83 LysaspleuarglsnglylsProvalgluvalProthrlyraspyrserglnhisthr 102
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 364 AGGTTACGACGACGCGGTGTCTACCTCGGACGCTGCTTGTGTTGAGGCACTTG 423
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 103 Argserlysgluhrillealpheasprplyasphalilelevalgluglylephe 122
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 424 GTGTTTACGACGACGAGATCCGAGACATGTTCCACTGCGCTTCTGTGACACGAC 483
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 123 AlaileuglnasnsanhrilleuargasprmetleAspvalylserlyvalAspthrAsp 142
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 484 TCCGAGTCAAGCTGTCTCGAAGATTCCTCGGAGCTGCGC--CGAGGAGGAGACTG 540
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 143 AlaaspleuargilleuargyleuthrArgasprthrlysgluarglyargthrmet 162
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 541 GACGACATTTCTGACGACATCACACCTTGTAAGCGCGCTTCCGAGGATTCGTGCTG 600
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 163 GluserValilleansnlyrleuasnsvalalargPromethsngluginpelleglu 182
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 601 CCGACAAAGAAGATGATGATGATATCCACGAGAGTGAACAATATGTTGCATC 660
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 183 ProthrlyserhisalalepillelleProgluglylyserhsnlyvalalalle 202
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 661 AACCTGATCGTGAACATCCGAGACATTTCTGAAT 696
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 203 Aspillemethrthrlyrsleuglnserleuvalser 214
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 4

US-09-107-532A-5024
 / Sequence 5024, Application US/09107532A
 / Patent No. 6585275
 / GENERAL INFORMATION:
 / APPLICANT: Lynn A Doucette-Stamm and David Bush
 / TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
 / ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
 / NUMBER OF SEQUENCES: 7310
 / CORRESPONDENCE ADDRESSES:
 / ADDRESSEE: GENOME THERAPEUTICS CORPORATION
 / STREET: 100 Beaver Street
 / CITY: Waltham

```

/ STATE: Massachusetts
/ COUNTRY: USA
/ ZIP: 02354
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: CD/ROM ISO9660
/ COMPUTER: PC
/ OPERATING SYSTEM: <Unknown>
/ SOFTWARE: ASCII
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/107,532A
/ FILING DATE: 30-Jun-1998
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 60/085,598
/ FILING DATE: 14 May 1998
/ APPLICATION NUMBER: 60/051571
/ FILING DATE: July 2, 1997
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Ariniello, Pamela Deneke
/ REGISTRATION NUMBER: 40,489
/ REFERENCE/DOCKET NUMBER: GTC-012
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (781)893-5007
/ TELEFAX: (781)893-8277
/ INFORMATION FOR SEQ ID NO: 5024:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 210 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ HYPOTHEICAL: YES
/ ORIGINAL SOURCE:
/ ORGANISM: Enterococcus faecium
/ FEATURE:
/ NAME/KEY: misc.feature
/ LOCATION: (B) LOCATION 1...210
/ SEQUENCE DESCRIPTION: SEQ ID NO: 5024:
/
US-09-107-532A-5024

Alignment Scores:
Pred. No.: 7.95e-36 Length: 210
Score: 434.00 Matches: 83
Percent Similarity: 62.33% Conservative: 51
Best Local Similarity: 38.60% Mismatches: 65
Query Match: 28.67% Indels: 16
DB: 4 Gaps: 4

US-09-896-522-3 (1-834) x US-09-107-532A-5024 (1-210)
QY 64 CGGCCCTTCTGATNAGGGTACGCGGCGGCACTGCCAGGGAAGTCAACGTTGTGAG 123
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 6 Lysproillellellyvalthrlyglyserglyserglyserlythrvalserarg 25
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 124 AAGATCATGAGATTTGCTGGAGACGAACGAGTGAACAGCGGACGAGGAGTGTGATC 183
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 26 Alailepheasnsanhrleproasphs-----Serleuemet 38
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 184 CTGAGCCAGACGAGCTTTACAAG-----GTCTGACGCGACGACGAGGACGCAAG 234
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 39 LeugluginasprserlyrlylsasprglnserhisleuserPheglugluargleuasn 58
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 235 GCCTTGAAGAAGACAGTGAATTTTGACCATCCAGATGCTTTGATATGATTTGATGCAC 294
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 59 Thr-----AsnlyrasphisprophealaPheasprasnleuile 73
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 295 AGGACTTGAGAGACATCGTGAAGGCAAAACGCTGAGAGTCCGACCTATGATTTGTG 354
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 74 GlnhsvalgllyAspleuenuasnlyrlysalalleglulysProvalylrasyrval 93
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 355 AACACTCAAGTTACCAAGACGCGGTGTCTACCTGCGGAGAGTGTGTTGTGAG 414
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 94 AlahtsthrargserglnalathrilleleinglnulysProlysgluvalilleleugln 113
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 415 GGCATCTGTGTTCTACGACGAGGAGATCCGAGACATGTTCCACTGCGCTTTCGTG 474

```

```

Db      114 GYIleuileuLeuGluAspGlnArgLeuThrGlySerLeuMetAspIleValIleVal 133
QY      475 GACACCGACTCCGACGTGAGCTGTCTCGAAGACTTCCGGACGTG---CGCCGAGG 531
Db      134 AspThrAspAspAspIleArgIleIleArgArgIleIleArgAspMetGluGluArgGly 153
QY      532 AGGAGCTCGAGCAGATTTCTGACCGCATACCACTTCCTCGAAGCGCGCTTCGAGAG 591
Db      154 ArgThrLeuAspSerValIleGluGlnThrLeuThrValValIleValIleValIleVal 173
QY      592 TTTCGCTCGCCGACAAAGATATGCGATGTGATTCATCCACGAGAGTGCACATATG 651
Db      174 PheIleGluProThrIleArgIleArgIleValIleValIleValIleValIleValIle 193
QY      652 GTTGCCATCAACTGATCTGTCGACACATCCAGACATCTTGAT 696
Db      194 ValAlaIleAspLeuIleThrThrIleValAlaSerPheLeuAsn 208

RESULT 5
US-09-198-452A-783
; Sequence 783, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Griffois, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198,452A
; NUMBER OF SEQ ID NOS: 1998-11-24
; SEQ ID NO 783
; LENGTH: 222
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-198-452A-783

Alignment Scores:
Pred. No.: 1,43e-33 Length: 222
Score: 412.00 Matches: 84
Percent Similarity: 63.01% Conservative: 54
Best Local Similarity: 38.36% Mismatches: 59
Query Match: 27.21% Indels: 22
Gaps: 6
DB: 4

US-09-896-522-3 (1-834) x US-09-198-452A-783 (1-222)
QY      73 CTGATAGGGGTGACGGCGGCGACCTGCGGAGAAAGTGCACCTGTGTGAGAAAGTATG 132
Db      8 ILeIleGlyIleThrGlySerGlyAlaGlyIleThrThrLeuThrGlnAsnIleLys 27
QY      133 GAGTGTGTCGAGCAAGAAAGAGTGCACGCGGACGCGAAGTGTATCTTCAGACG 192
Db      28 GluIlePheGlyGluAsp-----ValSerValIleCysGln 39
QY      193 GACAGGTTCTACAG-----GTCTGACGCGACAGCAAGAGCGCAAGCCTTGAAA 243
Db      40 AspAsnIleIleLysAspArgSerHisIleThrProGluGluArgAlaAsnIleLeu--- 58
QY      244 GGAACAGTACATTTGACCATTCAGATGCTTGTATATGATTTGATGACAGACTCTG 303
Db      59 -----TTPAspHisPheProAspAlaPheAspAsnLeuIleSerAspIle 74
QY      304 AAGAACATCGTGAAGGCAAAACGTGAGAGGTGCCGATATGATTTGTGACACTCA 363
Db      75 LysArgLeuLysAsnAsnGlnIleValGlnAlaProValPheAspPheValLeuGlyAsn 94
QY      364 AGG---TTACAGAGACACACGCTGTACCTTCGCGACGTGTTCTGTTGAGGCGATC 420
Db      95 ArgSerLysThrGlnIleGluThrIleLysProSerLysValIleLeuValGluGlyIle 114
QY      421 TTGTGTTTCTACACGACGAGATCCGGGACATGTTCCACCTGCGCCTTCTGTGACACC 480

```

```

Db      115 LeuValPheGluAsnGlnIleGluLeuArgAspLeuMetAspIleArgIlePheValAspThr 134
QY      481 GACTCCGACGTGACGCTGTCTCGAAGACTTCCGGACGTGCGC---CGAGGAGGAC 537
Db      135 AspAlaAspGluArgIleLeuArgArgMetValArgAspValGlnIleGlnIleAspSer 154
QY      538 CTGACACAGATTTGACGAGATACACACCTTCGTGAACCGCGCTTCGAGAGCTCTG 597
Db      155 ValAspCysIleMetSerArgIleLysSerMetValIleProMetHisGlnLysPheIle 174
QY      598 CTGCGCAACAAGATATGCTCGATGTATCTCCACGAGAGTGCACATATGTTGCC 657
Db      175 GluProThrArgIleArgIleValIleValIleValIleValIleValIleValIle 194
QY      658 ATCACTGTATCTGTCAG-----CACATCCAGACATTTGATGTGAC 702
Db      195 ThrAsnIleLeuSerGlnLysIleLysAsnHisIleGluAsnAlaLeuGluIleSerAsp 213

RESULT 6
US-09-252-991A-24671
; Sequence 24671, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; NUMBER OF SEQ ID NOS: 1998-07-27
; SEQ ID NO 24671
; LENGTH: 1150
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24671

Alignment Scores:
Pred. No.: 6.4e-08 Length: 1150
Score: 164.00 Matches: 70
Percent Similarity: 31.06% Conservative: 12
Best Local Similarity: 26.52% Mismatches: 108
Query Match: 10.60% Indels: 74
Gaps: 13
DB: 4

US-09-896-522-3 (1-834) x US-09-252-991A-24671 (1-1150)
QY      709 TGCAGATGTCACATTCAGATATCTCTGATGTCTGCACGATCAGATTGACACCA 650
Db      13 CysAlaThrArgAsnAlaSerCysValaCysTTPArg-----TTPArgPro 28
QY      649 TATTGTCCACTCTCTGGGATGATCATCATCGGCATCTTTTGTGCGACGACAACT 590
Db      29 ValAlaPro----- 31
QY      589 CTTGGAAGCGCGCTTCACGAAGTGTGTACTGCTCGAAGATTCCTGCTCCAGTCCCTCC 530
Db      32 -----ArgArgTTPSerSerAlaArgThrSerIleProCysProLys 45
QY      529 CTCGGCGCACGT-----CCGAGAACCTTTGAGACA 497
Db      46 AlaCysAlaArgAsnTTPHisAlaIleLysCysTTPValProIleCysCysSerSerSer 65
QY      496 GCTTACAGTTCGAGTGGGTGCCAGAGAGCGGACAGTGAACA----- 452
Db      66 GlyAlaAlaAsnAlaSerCysProArgArgAsnGlyArgArgThrProTTPProPro 85
QY      451 -----TGTCCCGATCTCTGCTGTGAAACA-----CCA 422

```

Db 86 AlaProThrThrcysargalacysproalaglyglyalalargThrSerThrglyala 105
QY 421 AGATGCCCTCAACAGAACACCGTCGCGAGGGTGAACCGCGTGTCTGTGTAACCTTG 362
Db 106 AlaAlaProaspThrAlaAlaProArgAsnAlaArgProThrTrpPaaenCysAlaProArg 125
QY 361 AGTGTG-----TCACAAATATCATAGTTCGGACCTCCACCGCTTTTGCCCT 317
Db 126 SerAlaGlyProTrpProProAlaSerAsnArgSerIleProThrPro-----Pro 143
QY 316 CCA-----CGATGTTCTTCAAGAGTCTGTGATCAAAATCATATTAACAAGGATCTGAT 263
Db 144 ProArgSerArgTrpThrProAlaSerAlaMetSerIlylProProAlaArgTrp--- 162
QY 262 GGTCAAAATGTACTGTCCTTTCAGAGGCTTGCGCTTCTGTCTGCGCGTCAAGACCTGT 203
Db 163 -----CysTrpCys-----ProTrpArgThrProTrpAlaAlaTrpSerSer 176
QY 202 AGAAGCTGTCTGCTGAGTGAACACCTTCGCGTCCGCTGTTCCACCTGCTTCTGTG 143
Db 177 ArgThrCysProAlaProAlaThrProThrArgThrIlylGlyAlaGlyArgIlyThr 196
QY 142 CCAAGCAATCCATGATTTTTCACACACGCTGCACTTCCGCTGGCAGTCCGCGCTCA 83
Db 197 ProArgArgCys---SerGluArgArgArgSerThrglyAlaCysAlaCys----- 212
QY 82 CCCTTACAGAGGCGCGTGTGCGGAGCGGTGCGGCTCCGCGCGGGGCTCTCGC--- 26
Db 213 -----SerAsnAlaAlaAlaGlyMetProArgArgProMetIleGluProArgAla 230
QY 25 AGTCTTGCCTC 14
Db 231 ThrLeuArgIleu 234

RESULT 7
US-09-199-637A-369
; Sequence 369, Application US/09199637A
; Patent No. 6355411
; GENERAL INFORMATION:
; APPLICANT: Ausubel, Frederick
; APPLICANT: Goodman, Howard M.
; APPLICANT: Rahme, Laurence G.
; APPLICANT: Mahajan-Miklos, Shalina
; APPLICANT: Tan, Man-Wah
; APPLICANT: Cao, Hui
; APPLICANT: Drenkard, Eliana
; APPLICANT: Tsongalis, John
; TITLE OF INVENTION: VIRULENCE-ASSOCIATED NUCLEIC ACID
; TITLE OF INVENTION: SEQUENCES AND USES THEREOF
; FILE REFERENCE: 00786/361002
; CURRENT APPLICATION NUMBER: US/09/199,637A
; CURRENT FILING DATE: 1998-11-25
; PRIOR APPLICATION NUMBER: 60/066,517
; PRIOR FILING DATE: 1997-11-25
; NUMBER OF SEQ ID NOS: 437
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 369
; LENGTH: 265
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; US-09-199-637A-369

Alignment Scores:
Pred. No.: 7.79e-08 Length: 265
Score: 160.00 Matches: 87
Percent Similarity: 35.62% Conservatives: 22
Best Local Similarity: 28.43% Mismatches: 107
Query Match: 10.34% Indels: 90
DB: 4 Gaps: 17

US-09-896-522-3 (1-834) x US-09-199-637A-369 (1-265)
QY 820 TGCTGAGCTCAATGTGACCGTTTGGCCAGAGGTGACATCCAGGAGGTGTCCTCGGCT 761

Db 11 CysArgSerProAsnValProArgCysCysArgSerGlyCys---GlyArgArgAlaAla 29
QY 760 CAGAAAAGTCCGCT-----TGAGTCCGCCAT----- 731
Db 30 GluSerArgSerIleProTrpProAlaProCysAlaAlaGlyAsnProProArgThrCys 49
QY 730 -----TGAGCCCTCCGCGTGCATTTGGAGTGCACCATTCAGAAATGCTGGATGT 677
Db 50 ArgGlySerSerAspLeuSerProValProArgAlaValProGlyGlnAlaAlaSerThr 69
QY 676 GCTGACAGATTCAGGTGATGAGCAACCATATTTGCCATCTCTCGTGGAGTATCATCCG 617
Db 70 ThrArgArgSerArgIleuProProAlaAlaProCys-----Arg 83
QY 616 CATATCTTTTGTGGCAGGAGCAACTCTCGAAGCCGCGCTTACGAAAGTGTACT 557
Db 84 HisGlySerArgHisAla---ArgSerProArgArgProIleuAspArgArg-----Arg 100
QY 556 GCGTCAGAAATCGCTCCAGGTCCTCCGCGGCGACAGTCCGAGAACTCTTCGAGACA 497
Db 101 AlaThrSerArgAlaPro---ProSerValHisAla----- 111
QY 496 GCGTCAGTCCGAGTCCGATGTCACACGAGGCGCAGTGAACATGTCGCGATCTCT 437
Db 112 -----ThrArgAspTrpArgIleValAlaGlyMetProAlaMetSerAlaPro 127
QY 436 GCGTGTGAACACCAAGATGCTCTCAACAGAACACAGTCCGAGGTGAGCACCGGTG 377
Db 128 Gly-----GluAspCysAlaHisTrp---ProArgProThrCysArgProPro--- 142
QY 376 TCTCTGTAACTGTAGTGTGTCAAAATCATAGTGGGACCTTCACCGTTTGGCT 317
Db 143 -----AlaGlySerLeuArgProAlaArgAlaProAlaArgPro 156
QY 316 CCAAGATGTTCTTCAAGTCTGTGATCAAAATCATTAACAAGCATCTGATGTC 257
Db 157 ProArgSerAlaAlaLeuProSerThrAla-----GlyCys 167
QY 256 AATTGACTGTCTTTCAGAGGCTTGCGCTTCTGTCTGCGC-----TAGGA 209
Db 168 GlyCysVal-----ProThrAlaValProArgSerGlyArgProGly 181
QY 208 CCTTGTGAAC----- 197
Db 182 ProAlaSerThrclyGlnArgIlylProArgAspArgAlaArgSerArgAlaAsn 201
QY 196 ---TGCTGTGCTCAGATGACCACTTCGCGTCCGCTGTTCCACCTGTTCTGCCA 140
Db 202 AlaAlaProGlyProGlyArgSerProAlaThrSerAla---ProGlyArgSerProPro 220
QY 139 GCAATCTCATGATTTTTCACACACGCTGACTTCCGCTGGCAGTCCGCGCTCACCC 80
Db 221 GlyCys-----SerAlaThrAspArgAlaAlaArgArgSerArg 233
QY 79 CTATCAGAAAGGCGCGCTGTGTGCGACGTCGCGCTCCGCGCGGCGCTCTCGCAATCT 20
Db 234 LeuProProArgIleProHisAlaHisSerArgProAlaProProGlySerProAlaVal 253
QY 19 CGCCTCCCGCCGAGCCA 2
Db 254 ArgArgAlaProAlaPro 259

RESULT 8
US-09-252-991A-25569
; Sequence 25569, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A

```
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 25569
; LENGTH: 274
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25569

Alignment Scores:
Pred. No.: 1.13e-07 Length: 274
Score: 158.50 Matches: 72
Percent Similarity: 35.50% Conservative: 21
Best Local Similarity: 27.48% Mismatches: 78
Query Match: 10.25% Indels: 91
DB: Gaps: 16

US-09-896-522-3 (1-834) x US-09-252-991A-25569 (1-274)
QY 763 GCTCAGAAAAGGTCCTGTTG-----AGCTCCGCCATTGGACCTTCCTCGTGCATT 710
Db 88 AlaArgArgArgSerSerValArgArgAmArgGlyHisThrArgArgValAlaThrAla 107
QY 709 TGCAGATGTCACCATTCAGATGTCCT-----GGATGCTGTCACATCAGTGTGA 659
Db 108 ProArgCysArgArgArgArgGlyThrProAlaAlaSerGlyAsnProAlaArgAla----- 125
QY 658 TGGCAACCATATGTCACATCTCTCGTGGAGTATCACATCGCATACTTCTTTGTCGCA 599
Db 126 -----ArgHisThrAla----- 129
QY 598 GCGCAAACTCTCGAAGCGCGCTTCACGAAGTGTTACT---GGTCAGAATGCT 542
Db 130 -----ProArgArgArgAlaArgArgGlyCysAlaArgAlaSerArgGly 146
QY 541 CCAAGTCCTCCCTCGGGCGACGTCGCGAAGACTCTTCGAGACGCTGACGTCGAGT 482
Db 147 ProGlyGly----- 149
QY 481 CCGTGTCCAGAGAGCGCAGGTGA-----ACATGTCGCGATCTCTGCTGTAGA 428
Db 150 -----ProArgArgAlaAlaArgAlaProValAlaAlaProGlySerProGly--Arg 166
QY 427 ACACCAAGATGCTCTCAACGAACACAGTCGCGAGGTAGACCACTGCTGTGTA 368
Db 167 SerAlaThrAlaArgProSerAlaProArgAlaProAlaThrProPro----- 182
QY 367 ACCTTGAGTGTGTCACAAATCATAGTGGCAGCTCCACCGTTTGGCCCTCCACATGT 308
Db 183 -----AspAlaProProAlaArgProAlaAlaCys 192
QY 307 TCTTCAGAGTCCTGTGATCAATCATATTATCAAGATCGATGTGATCAAAATGTACT 248
Db 193 AlaAlaSerGlyCysArgSerTrpArgTrpArgTrpArgTrpAlaGlyAlaVal 212
QY 247 GTCTTCAGAGCTTGAGCTTCTGCTGCTGCGCAGACCTGTGAACTGTCTGGC 188
Db 213 ArgAlaTyArgPro-----LeuSerArgProProCys---ProCysAlaGly 227
QY 187 TCGAGTAGACACCTTCGCTGCGCGCTGTTCACCTGTTCTGTCGACGAATCCATGA 128
Db 228 AlaGly-----SerAlaGlyAlaCysArgSerArg--LeuPro----- 239
QY 127 TCTTTCACACACGCTGACTCCCGCTGGCAGTGGCGC---CGCTACCCCTATCAGA 71
Db 240 -----ArgCysArgProGlySerProSerSerGly 249
QY 70 AGGCGCCCTGAGTCGACGCTCGGCTCCGCGCGGAGCTTCGACATCTTCGCTCCG 11
Db 250 ArgSerCys-----ArgProProAlaSerProGlyGluSerProCysIlePro 265
```

```
QY 10 CCGAG 5
Db 266 AlaArg 267

RESULT 9
US-09-252-991A-26899
; Sequence 26899, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 26899
; LENGTH: 247
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-26899

Alignment Scores:
Pred. No.: 1.35e-07 Length: 247
Score: 157.50 Matches: 83
Percent Similarity: 34.60% Conservative: 17
Best Local Similarity: 28.72% Mismatches: 84
Query Match: 10.18% Indels: 105
DB: Gaps: 21

US-09-896-522-3 (1-834) x US-09-252-991A-26899 (1-247)
QY 688 TGTTCAGATGTCCTGACGATCAGATGATGTCACCATATGTCACATCTCTCGTGGCA 629
Db 2 CysProSerCysSerProProSer-----ProCysSerProLeuAlaArgTrp 17
QY 628 TGATCATGCGCATCTTCTTTGCGCAGCAGACATCTCT-----CGAAG 581
Db 18 ProSerProThrProAlaThrAlaAsnThrAlaTrpSerProAlaSerProIleArgSer 37
QY 580 CCGGCTTCACGAAGTGTGTTACTGCGTCAGAAATCTGCTCCAGGTCCCTCCGCGCA 521
Db 38 ProAlaSerThrThrCysTrpArgCysSerProSerAlaSerGlyProProSerAsnArg 57
QY 520 CGTCCGAGAACTCTTCGACAGCGCTGACGTCGCGAGTGGGTGCGAGAAAG----- 467
Db 58 GluPro-----ArgValSerProCysProAlaProSerSer 69
QY 466 -----GGCCAGGTGAGACATGTCGCGATCTCTGCTGTAGAAC----- 425
Db 70 AlaProCysTrpSerAlaGlyCysSerAlaSerArgAlaCysAsnCysArgSerTrpIle 89
QY 424 -----CCAAGATGC----- 416
Db 90 ProAlaSerProProArgCysSerProSerAlaSerAlaTrpProSerProCysAlaArg 109
QY 415 -----CCTCAAGACGAACGTCGCGAGGTAGAC----- 383
Db 110 ArgCysArgTrpArgTrpProProProArgCysSerProTrpProThrGlySerProThr 129
QY 382 CCGTGTCTCTGTAACCTTGAAGTGTGTCACAAATCATAGTGGCAGCTCCACCGTTT 323
Db 130 AlaTrpSerCys-----ArgThrSerProAlaProGly 140
QY 322 TGC-----CTTCACGATGTTCTTCAGATTC----- 296
Db 141 CysThrArgSerValSerSerProProProProArgCysThrProProAlaThrArgTrp 160
```

```

QY 295 TGTGATCAATATCATATCAAGGATGTGGATGTGCAAAATTGATCTGCTTTCAAG 236
    |||||
Db 161 CysAlaSer-----TrrArgArgArg-----ProArg 169
QY 235 CCTGGCCCTTCTGCTGTGCGTCAGACCTGTGAGAACCTGCTGCTGAGATGACCA 176
    |||||
Db 170 ProTrrP---TyrValSerProGlyTrrPro---ArgArgAlaProGlySerGlyCysTrr 187
QY 175 CCTTCGCGTCCGCTGTTCACACCTGCTGCTGCTCCAGCAACTCATATCTTCTCACACA 116
    |||||
Db 188 Pro-----AlaGluProProArg---LeuPro-----GlnArgLeu 198
QY 115 CGGTGCACTTCCGCG---TGACAGTGCCTGCGCTCACCC---CTATCAGAAAGGCGCT 62
    |||||
Db 199 TlSerThrArgArgProTrrPrrCysThrArgArgArgProTrrValCysAlaSerTleAla 218
QY 61 GGTGCGGACGCTGCG-----CCTCGCGCGCG----- 35
Db 219 ArgAlaSerGlyArgSerSerThrGlyValProGluThrArgSerAlaSerSerArg 238
QY 34 GCGTCTCGCAGTCTTCCGCTCCGCGCG 8
Db 239 GlySerHisSerMetLeuSerProPro 247

RESULT 10
US-09-252-991A-31116
; Sequence 31116, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 31142
; SEQ ID NO 31116
; LENGTH: 664
; TYPE: PRN
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-31116

Alignment Scores:
Pred. No.: 2,51e-07 Length: 664
Score: 157.00 Matches: 90
Percent Similarity: 29.09% Conservatve: 15
Best Local Similarity: 24.93% Mismatches: 105
Query Match: 10.37% Indels: 151
DB: 4 Gaps: 16

US-09-896-522-3 (1-834) x US-09-252-991A-31116 (1-664)
QY 15 AGCGGAAGATCGGAGAGCCC-----CGCGCCGAGAGCGGACCGCTCCGACACGCG 65
    |||||
Db 265 ArgArgArgArgAlaSerProTrrPrrArgArgArgArgGlyArgProGlyGlnGlyAla 284
QY 66 GCCCTTCTATAGGGGTGAGCGCGGCGGACCTGCGACGAGGGAAGTCGACCGCTGTGTGAA 125
    |||||
Db 285 Ala-----HisArgArgGlyHisArgGlyArgProArgSerGlyAla 298
QY 126 GATCATGAGATT-----GCTGAGACAGAAACGAGGTGGAACAGCG 164
    |||||
Db 299 ArgHisGlyValHisAlaSerArgArgGluGlnProGlySerGlyArgProGlyArgPro 318
QY 165 GCGACGGAAGGTGTCTCTGAGCCGAGACAGGTTTACAGGT----- 209
Db 319 AlaProGlyArgProArgProAlaGlnGlyArgGlnGlyArgHisArgLeuArg 338
QY 210 -----CCTGACGCGAGAGCAAGAGCGCCTTGAAGAGACATGATTTTGA 260

```

```

Db 339 ProHisArgProGlyArgGlnGlyArgGlnHisGly----- 350
QY 261 CCAATCCAGATGCTTTGATTAATGATTTGATGCAACAGAACTGTGAAGACATGTCGAGG 320
    |||||
Db 351 -----GlnGlnGluProProGlyGly 356
QY 321 CAAGAAGGTGAAGTGGCGGACCGATATGTTTTGTGACACATCAAGTTACAGAGACAC 380
    |||||
Db 357 ArgArgGlyAlaGly-----AspArg 363
QY 381 GGTGTCTACCTCTGCGGAGCGTGTCTGTGTGAGG----- 416
    |||||
Db 364 GlyAlaLeuProValProArgGlnHisProArgGlyValAlaSerHisArgHisHisPro 383
QY 417 -----CATGTT-----GGTGTTCACACCAAGGA 440
    |||||
Db 384 AlaValAspAlaValHisSerHisArgSerGlyArgGlnGlnGlyGlnArgGlnProArg 403
QY 441 GATCCGGACATGTTCCACCTGGCGGCTTTCGTGACACCGGACGTCGAGCTGTC 500
    |||||
Db 404 GluProArgArg-----AlaGlyLeuArgHisHisArgArgArgArgGlyArg 420
QY 501 TCGAAGAGT---TCTCCGAGACGTCGCGGAGGAGGGA----- 536
    |||||
Db 421 ArgArgGluArgHisProProProGlyProArgProGlyProProArgProProAlaArg 440
QY 537 -----CCTGAGCAGAT 548
    |||||
Db 441 ProArgGlyThrProArgGlyLeuArgArgValGlnGlyGlyAlaProGlyAlaGly 460
QY 549 TCGAGCGAGTACACACCTGCTGTAAGCGGCTTCGAGAGTTGCTGCGGACAA 608
    |||||
Db 461 LeuArgProArgHisHis-----HisGlyGlyValProAlaAspLeu 474
QY 609 GAAGTATGCCAGTGTATCATCCACGAGAGTGCACATAT-----GGTTGCCAT 659
    |||||
Db 475 Arg-----ProHisArgGlyArgGlyGluAspValProProArgGlyValHis 490
QY 660 CAACSTGATGTGACGACGACATCCAGGACATTTGAATGTGCATC----- 705
    |||||
Db 491 ArgGlyHisArgGlyAlaArg---ArgAspArgProLeuGlyAspLeuArgProGlyGlyAs 510
QY 706 -----TGCAAATGCGACCGGAGGCTCAATGGCGGAGCTAC----- 744
    |||||
Db 510 ArgAlaValHisHisArgGlnGlyGlnGlyGlyGluLeuArgHisAlaProArgAl 530
QY 745 -----AA 746
    |||||
Db 530 AsArgGlyLeuArgThrGlyAlaAlaLeuGlyAlaArgAlaProArgAlaGlyGlyArgArg 550
QY 747 GCGGACCTTTTTCGACGCGGAGGACCGGAGATGCTGACCTTGGCAACGATGAC 806
    |||||
Db 550 gArgProArgArgAspProAlHisArgAlaGlyGlyLeuAlaAspGlyGlnArgValHis 570
QY 807 T 807
Db 570 s 570

RESULT 11
US-09-252-991A-24899
; Sequence 24899, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190

```


Qy	502	GAGACAGCCTGACGTCGGAGTCGGTGTCTCCAGAAAGAGCGAGGTGAAACATGTCGGCA	443
Db	122	AlaThrAlaArgArgArgAlaThrGlyArgAlaGlyArgGlyAlaGlyGly-----GlyGly	139
Qy	442	TCTCTGGCTGTGAAACACCAAGATGCGCTTCAACAGAAACACGTCGCGAGGTAGACCA	383
Db	140	ArgGlyGlyGlyArgArgProAlaThrProAspLeuHisLeuLeuProGlyAlaAsp	159
Qy	382	CGGTGGCTCTGGTAACTTGAATGTGTCTCAAAATCATAGTGGCA-----CCT	332
Db	160	ProArgGlyThrGlyGlyAlaAspProAlaArgGlyLeuLeuArgProAlaHisArgGlyAsp	179
Qy	331	CCACGCTTTGGCCCTCCACGATGTTCTTCAAGTCTCTGCATCAATCATTAACAAG	272
Db	180	ArgProArgLeuProArgArg-----AlaGlnArgHisArgProThr	193
Qy	271	CATCTGGATGGTCAAAATTTGATCTGCTTCAAGGCTTGGCTTGGCTGCGCGTCA	212
Db	194	HisArgAlaGlyGlnIle-----GlnAspProArgGlyProAspSerLeuProGly	210
Qy	211	GGACCTTTGTAGAACCTGTCGAGTCAGGATACCAACCTTCCGCTGCGCTTCCACCT	152
Db	211	-----ThrsProGly--GlyThrProArgAlaAlaGlyGlnArgProAla	225
Qy	151	CGTCTGTCCAGCAACTCCATGATCT-----TCTCAC	119
Db	226	SerAspLeuProGlyValGlnArgGlyLeuLeuGlyLeuHisArgArgGlnProAspPro	245
Qy	118	ACACGGTGCACCTTCCGCTGGCAGTGCAGCGCGCTCAACCCATATCAGAAAGGCCGTGGT	59
Db	246	ArgArgSerGlnArgArgGlyAspProArgArgProProAlaGlyArgAlaAlaAlaArg	265
Qy	58	GCGGACGGTGGCGCTCGGCGGCGGGGCTTGCACATCTTCCGCTC	14
Db	266	ProArgGlyProArgProAlaArgProAspArgValAlaAlaArgLeu	280

```

RESULT 13
US-09-252-991A-21436
Sequence 21436, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERGININOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 21436
LENGTH: 265
TYPE: PRF
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-21436

Alignment Scores:
Pred. No.:      2.24e-07      Length:      265
Score:          155.50      Matches:      84
Percent Similarity: 31.86%      Conservative: 17
Best Local Similarity: 26.50%      Mismatches:  99
Query Match:    10.05%      Indels:      117
DB:             4          Gaps:         17

US-09-896-522-3 (1-834) x US-09-252-991A-21436 (1-265)
QY      775 GGTGATCCCTGGCTGACGAAAGAGTCCGCTTGAGTCGCCGCAATGGAGCCCTCCTCGGT 716
      ||| |||
      ||| |||
Db      1 GlypHePrOmAlaPrOxArgYsSerAlaSerYsArgSerPrOxan-----CGATCGAGT 662
QY      715 GCCATTTGCAGATGTCAACCATTCAGAAATGCTCGATGTGCTGCA-----CGATCGAGT 662

```

Db 16 ---ValProAlaGlyCysValArgSerGlyCysGlyAlaArgAlaAlaGlySerAla 34
QY 661 TGATGGCAACCATATTGTCCACTCTCCGTGGGATGATACATCGGCATCTTTGTGCG 602
Db 35 ProThrProAlaProCys-----Ala 41
QY 601 GCAGGCGCAACTCTCCGAAAGCGCGCTTCAACGAAGTGGTGTACTGGCTCAAACTTGCT 542
Db 42 AlaGlyAsnProProAlaGlyThr-----CysValArgGlySer---SerAla 54
QY 541 CCAAGTCCCTCCCTCCGGCGACCGTCCGGAGAACTCTTCAGACAGCGCTCAACCGGAGT 482
Db 55 LeuSerProValProAlaGlyAlaValProGlyGlnAlaAlaSerThrThr---ArgAlaSer 73
QY 481 CGGTGTCCA----- 473
Db 74 ArgAlaArgProProProAlaAlaProCysAlaGlnIsoleuSerAlaAlaGlySerProAla 93
QY 472 -----CGAAGAGCGCGAGGTGAACATGTCC----- 446
Db 94 ArgAlaLeuAspArgAlaGlyAlaThrSerArgAlaProProSerAlaIleAlaValAlaG 113
QY 445 -----GGATCTCTCGGCTGTGAACACCA-----AAGTCCCTCA 410
Db 114 AspThrArgIleValAlaGlyMetProAlaMetSerAlaProGlyGluAspCysAlaIle 133
QY 409 ACAGAACACGACGCGGAGGTAGACACACCGTGGTGTAACTTGAAGTGTGTACAA 350
Db 134 Trp---ProAlaArgProThrCysValArgProPro-----AlaGlySer 145
QY 349 AATCATAGTGTGGACACTCCACCGTTTGTGCTCCAGATGTTCTTCAAGTCCCTGTGA 290
Db 146 ArgAlaArgProAlaArgAlaProProAlaGlyArgProProArgSerAlaLeuProSerThrAla 165
QY 289 TCAATATCATTAACAAGCATGTGAAGTGTCAAAATGTACTGTGCTTTCAAGCGCTTG 230
Db 166 -----GlyCysAlaCysVal----- 170
QY 229 CCTTGTGCTGTGCCG-----TCAGACCTTGTGAAC----- 197
Db 171 ProThrAlaValProAlaGlySerGlyAlaGlyProGlyProAlaSerThrGlyGlnAlaGly 190
QY 196 -----TGTCTGTGCTCAGATGACCACT 173
Db 191 ProAlaArgAspArgAlaArgSerArgAlaAlaAsnAlaAlaProGlyProGlyAlaGlySerPro 210
QY 172 TCGCGTCGCGCTGTTCACCTCGTTCTGTGTCCAGCAACTCCATGATCTTCTACACACGG 113
Db 211 AlaThrSerAla---ProGlyArgSerProProGlyCys----- 222
QY 112 TCGACTTCGCGCTGCGAGTCCGCGCTTCAACCCCTAACAAGAGCGCGCTGTGCGGAC 53
Db 223 SerAlaThrAspArgAlaAlaAlaArgArgSerArgLeuProProAlaGlyMetProIleAlaIle 242
QY 52 GATGCGGCTCCGCGCGCGGCTCTCGAGTCTTTCGACTCCCGCGGAAGCA 2
Db 243 SerArgProAlaProProGlySerProAlaValAlaArgAlaProAlaPro 259
RESULT 14
US-09-252-991A-30843
; Sequence 30843, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190

```

Db      166 -----GlyCysAlaCysVal----- 170
Qy      229 CCTTCTGCTGTGCCG-----TCAGACCTTGTGAAC----- 197
      |||:::|||||
Db      171 ProthrAlaValProArgSerGlyArgProGlyProAlaSerThrGlyArgTyrGly 190
      |||:::|||||
Qy      196 -----TGTCTGTGCTCAGATGACACCT 173
      |||:::|||||
Db      191 ProArgSparGlyAlaArgSerArgArgAlaAsnAlaAlaProGlyArgSerPro 210
      |||:::|||||
Qy      172 TCAGCTGCGCGGTGTTCACCTGTTCTGTGCCAGAACCTCAGATGATCTTCACACACGG 113
      :::::|||||
Db      211 AlaThrSerAla--ProGlyArgSerProProGlyCys----- 222
      :::::|||||
Qy      112 TCAGCTTCCGCGCTGCAGTCGCGCGCTTCACCCCTATCAGAAAGGCGCGTGTGCCGAG 53
      |||:::|||||
Db      223 SerAlaThrSparGlyAlaAlaArgArgSerArgLeuProProAlaMetProHisAlaHis 242
      |||:::|||||
Qy      52 GGTGCGGCTCCGCGCGCGGCTCTCGCAGTCTTGCGCTCCGCGGAGCGCA 2
      |||:::|||||
Db      243 SerArgProAlaProProGlySerProAlaValArgArgAlaProAlaPro 259
      |||:::|||||

RESULT 14
US-09-252-991A-30843
; Sequence 30843, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS.
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190

```

```

; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 30843
; LENGTH: 663
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-30843

Alignment Scores:
Pred. No.: 6.41e-07      Length: 663
Score: 153.00           Matches: 102
Percent Similarity: 27.88%      Conservative: 31
Best Local Similarity: 21.38%      Mismatches: 114
Query Match: 9.89%           Indels: 230
DB: 4                     Gaps: 26

US-09-896-522-3 (1-834) x US-09-252-991A-30843 (1-663)

QY 832 AGTGGGCTGCTGCTGACTCCAAATGTGACCGTTTGCAGAGGTCCAGATCCCAAGGT 773
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 75 AlaGlyAlaCysCysTrpLysThrArg---AsnValAlaGlyArgSerAlaThrAspAla 93
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 773 ----- 773
Db 94 TrpSerSerAlaProAlaProAlaAsnCysCysLeuPheProAlaGlyProSerPro 113
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 772 -----GGTCCCGCTCAGAAA-----AGTCCGCTGT-----AGC 740
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 114 AlaSerProAlaProAlaAlaCysArgArgTrpProAlaGlyAlaCysHisTrpProAla 133
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 739 TCCGCCATTGGAACCTCTCGTGCCTATTGC----- 707
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 134 SerAlaTrpTrpLysLeuAlaProAlaArgCysCysTrpProAlaProProAlaProAla 153
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 707 ----- 707
Db 154 SerAlaGlyArgAlaCysCysAlaSerProAsnArgArgGluProTrpProPro 173
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 706 -----AGATGTCAACCATTCAGAAATGT-----CCTGATGT----- 677
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 174 SerProTrpAlaSerArgAlaGlyProAlaSerCysGlyArgProProAlaCysSerPro 193
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 676 -----GCTGCACGATCAGATTGATGGCAA 653
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 194 ValAlaThrAlaProThrAlaThrCysSerProProSerAlaArgSerAla---TrpLys 212
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 652 CCA---TATTGTCCACTCTCTGTGGATGATCAATCGCATTTCTTTTGTCCGACGCC 596
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 213 ProCysAlaCysAlaLysAlaValGly-----SerAla 223
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 595 AGAATCTCTCGAAGG-----CCGACTCAGCAAGGTGTGATCGCTCAGAAATCTGCT 542
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 224 ArgSerProAlaAsnAlaTrpProAlaAlaSerAlaTrpCysProThrCysAsnSer--- 242
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 541 CCAAGTCTCTCTCTCGGCGACGTCCTCCGAGAACTTTGAGACAGCTGACGTGAGT 482
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 243 -----AlaArgProSerAlaIleAlaSerThrAla---ArgArgSer 255
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 481 -----CGTGTCCACGAAGAGCGCAGGATGTCCTCC 446
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 256 ProSerThrAsnGlyArgProAlaAlaCysProThrThrThrProProAlaAsnAlaPro 275
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 445 GGATCTCTCTGCTGAGAACCA----- 422
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 276 AlaSerAlaAlaAlaAsnTrpProTrpTrpLysAlaProSerProAlaMetProProSer 295
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 421 -----AGATGCCCTCAA 410
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 296 ThrSerAlaThrArgLeuAlaAlaCysGlyHisAspGlyAspAlaGlyArgAlaSerGln 315
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 409 ACAGAACCAAGCTCCG-----CAGCGTAGACCA 383
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 316 ThrProTrpArgAspThrSerProCysValProAsnCysValaAsnTrpArgArgArgIle 335
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

QY 382 CCGNGATCTCGGTAACCTTGAAGTGTACAAATCATGATGTCGGACCTCCACGGTTT 323
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 336 ProTrpSerAlaValAla-----LysThrSerProTrpProArg 348
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 322 TGCCCT-----CCACGATGTTCTTCA----- 302
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 349 TrpProGlyMetProAlaGlyProArgProSerCysThrAlaValAlaAlaTrpAlaPro 368
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 301 -----GATCTCTGTGCATCAATCATTTATCAAGG 272
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 369 AlaArgGlyGlyTyraAlaAlaAlaProAlaArgAsnSerCysSerAlaGlyArgLeuArgArg 388
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 271 CATCTGATGTCAAAATTTGACTGTCTTCAAGGCTTGCGCTTCT-----GCTCTG 218
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 389 ArgVal---ArgArgSerAlaAlaArgAlaTrpArgProTrpProAlaGlyArgAlaThr 407
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 217 CCGTCAGGACCTTGAGA----- 200
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 408 ProAlaSerAspAlaArgArgAspSerAlaSerArgGlyArgProAlaThrAlaAlaAsp 427
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 199 -----ACC 197
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 428 HisProAlaAlaTrpValSerAlaAlaAlaArgArgThrSerSerAlaProIleAlaGlySer 447
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 196 TGTCTGAGTCAGAGATGACCACTTCGCGCTGTTCCACTGTTCTGTCCAGCA 137
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 448 AlaProGlySerGlyThrAlaProAlaArgCysHis-----ProAlaArgLysArgGlyAla 465
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 136 ACTCGATGATCTTCTCACAACGCGTCACTTCCCGCTGCAGAGTC----- 92
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 466 GlyPro---AlaSerThrGlyArgSerArgArgTrpAlaAlaAlaProAlaArgArg 484
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 91 CGCCGTCACCCCTATCAGAGAAGCGCTGTGCGGACGTCGCGCTCGCGCGCGGCG 32
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 485 AlaArgThrGlyIleArgArgThrAlaAlaAlaAlaCysTrpProProAlaProGly 504
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 31 -----TCTCGAGTCTTGCGCTCCCGCGCAAGCA 2
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 505 ArgArgThrProAlaAlaAlaArgArgSerAlaArgThrAlaProArgPro 521
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 15
US-09-252-991A-24055
; Sequence 24055, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 24055
; LENGTH: 582
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24055

Alignment Scores:
Pred. No.: 6.75e-07      Length: 582
Score: 152.50           Matches: 89
Percent Similarity: 30.94%      Conservative: 23
Best Local Similarity: 24.59%      Mismatches: 111
Query Match: 10.07%           Indels: 139
DB: 4                     Gaps: 18

US-09-896-522-3 (1-834) x US-09-252-991A-24055 (1-582)

```


This Page Blank (uspto)

Result No.	Score	Query Match	Length	DB	ID	Description
1	1450	95.8	277	22	AAM39502	Human polypeptide
2	1450	95.8	277	22	AAB93941	Human protein sequ
3	1450	95.8	277	23	AAE16592	Human 57658 protei
4	1450	95.8	296	22	AAE41288	Human polypeptide
5	1372.5	90.7	276	22	AAE65056	Human uridine kinase
6	1363	90.0	260	22	AAO14412	Protein of a human
7	1325	87.5	277	22	AAO14413	Protein relating t
8	951	62.8	190	23	AAB89353	Human polypeptide
9	934	61.7	261	22	AAE38694	Human polypeptide
10	934	61.7	261	22	AAB73493	Human transferase
11	902.5	59.6	337	23	ABP41393	Human ovarian anti
12	897.5	59.3	337	23	AAB56582	Human prostate can
13	889	58.7	335	22	AAE40480	Human polypeptide
14	788	48.7	260	22	AAB662307	Drosophila melanog
15	589	38.9	326	22	ABG27216	Novel human diagen
16	579	38.2	120	23	ABG70328	Human MBDT protein
17	578	38.2	481	21	AAE34091	Zea mays protein f
18	578	38.2	490	21	AAE34090	Zea mays protein f
19	578	38.2	512	21	AAE34089	Zea mays protein f
20	570.5	37.7	500	21	AAE47030	Arabidopsis thalia
21	570.5	37.7	511	21	AAE47029	Arabidopsis thalia
22	545.5	36.0	548	23	AAU99000	Human uridine kinase
23	542	35.8	614	22	AAB69801	Drosophila melanog
24	542	35.8	614	22	AAB67004	Drosophila melanog
25	540.5	35.7	548	22	AAB97433	Novel human protei
26	540.5	35.7	578	22	AAU28181	Novel human secre
27	529	34.9	415	21	AAE51786	Arabidopsis thalia
28	529	34.9	433	21	AAE51785	Arabidopsis thalia
29	529	34.9	441	21	AAE51784	Arabidopsis thalia
30	529	34.9	443	21	AAE51749	Arabidopsis thalia
31	529	34.9	461	21	AAE51748	Arabidopsis thalia
32	529	34.9	469	21	AAE51747	Arabidopsis thalia
33	526.5	34.8	540	21	AAE41657	Human ORF4221
34	518	34.2	477	21	AAE45808	Arabidopsis thalia
35	518	34.2	500	21	AAE45807	Arabidopsis thalia
36	517.5	34.2	466	21	AAE48809	Arabidopsis thalia
37	507.5	33.5	496	21	AAE48426	Arabidopsis thalia
38	507.5	33.5	498	21	AAE48425	Arabidopsis thalia
39	507.5	33.5	521	21	AAE48425	Arabidopsis thalia
40	507.5	33.5	521	21	AAE48425	Arabidopsis thalia
41	507	33.5	465	21	AAE48427	Arabidopsis thalia
42	507	33.5	487	21	AAE48427	Arabidopsis thalia
43	490.5	32.4	274	21	AAE52640	Arabidopsis thalia
44	490.5	32.4	292	21	AAE52639	Arabidopsis thalia
45	490.5	32.4	300	21	AAE52638	Arabidopsis thalia

```
XX
PN      NO200153312-A1.
XX
PD      26-JUL-2001.
XX
PF      26-DEC-2000; 2000WO-US34263.
XX
PR      21-JAN-2000; 2000US-0488725.
PR      25-APR-2000; 2000US-0552317.
PR      09-JUL-2000; 2000US-0598042.
PR      19-JUL-2000; 2000US-0620312.
PR      03-AUG-2000; 2000US-0653450.
PR      14-SEP-2000; 2000US-0662191.
PR      19-OCT-2000; 2000US-0693036.
PR      29-NOV-2000; 2000US-0727344.
XX
XX      (HYSE-) HYSEQ INC.
XX
PI      Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI      Wang J, Wang Z, Mehran T, Xu C, Xue AJ, Yang Y, Zhang J;
PI      Zhao QA, Zhou F, Goodrich R, Drmanac RT;
XX
XX      WPI; 2001-442253/47.
DR      N-PSDB; A158658.
XX
PT      Novel nucleic acids and polypeptides, useful for treating disorders
PT      such as central nervous system injuries -
XX
PS      Example 4; SEQ ID NO 2647; 10078bp; English.
XX
XX      The invention relates to human nucleic acids (A157798-A161369) and
XX      the encoded polypeptides (AAM3642-AAM42213) with nootropic,
XX      immunosuppressant and cytostatic activity. The polynucleotides are useful
XX      in gene therapy. A composition containing a polypeptide or polynucleotide
XX      of the invention may be used to treat diseases of the peripheral nervous
XX      system, such as peripheral nervous injuries, peripheral neuropathy and
XX      localized neuropathies and central nervous system diseases, such as
XX      Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
XX      lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
XX      utilisation of the activities such as: Immune system suppression,
XX      CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
XX      and thrombolytic activity, cancer diagnosis and therapy, drug screening,
XX      CC assays for receptor activity, arthritis and inflammation, leukaemias and
XX      CC Note: The sequence data for this patent did not form part of the printed
XX      CC specification.
XX
SQ      Sequence 277 AA;

Alignment Scores:
Pred. No.:      1.5e-112      Length:      277
Score:          1450.00      Matches:      277
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      95.77%      Indels:      0
DB:              22          Gaps:      0

US-09-896-522-3 (1-834) x AAM39502 (1-277)
QY      1 ATGGCTTCGGCGGAGCGAAGACTGCGAGAGCCCGCGCGAGGCGGACCGTCCGAC 60
      |||||
DB      1 MetAlaSerHisIaGIyGIuAaPcYgIuSerProIaProGIuAlaAaPaRgProHis 20
QY      61 CAGCGGCCCTCTCTGATAGGGGGTGAAGCGCGGCACTGCGACGGGAAGTGCACCTGTGT 120
      |||||
DB      21 GlnArgProPheLeuIleGIyValSerGIyGIyThrIaSerGIyLysSerThrValCys 40
QY      121 GAGAGATCATGAGTGTCTGGGACAGAGAGGTGAAGACGGCGGACGGGAGGTGCTC 180
      |||||
DB      41 GlnIysIleMetGluLeuLeuGIyGlnAsnGIuValGIuGlnArgGlnArgLysValVal 60
QY      181 ATCCTGAGCCAGACAGGTTCTTAAGAAGTCTGACGCGACAGACGAGGCAAGGCTTG 240
      |||||
```

```
DB      61 IleLeuSerGlnAspArgPheTyrLysValLeuThrAlaGluGlnLysAlaLysAlaLeu 80
QY      241 AAAGACAGTACATTTTGAACCATCCAGATGCGCTTTGATTAATGATTGATGCACAGACT 300
      |||||
DB      81 LysGIyGlnTyrAsnPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThr 100
QY      301 CTGAAGACATCGTGGAGGGCAAAACGGTGGAGGTGCGACCCATATGATTTTGGACACAC 360
      |||||
DB      101 LeuLysAsnIleValGluGluLysThrValGluValProThrTyrAspPheValThrHis 120
QY      361 TCAAGGTTACCAAGACCAACGGTGTCTACCCCTGCGAGCGTGGTCTGTTTGAAGGCATC 420
      |||||
DB      121 SerArgLeuProGIuThrThrValValTyrProAlaAspValValLeuPheGluGlyLe 140
QY      421 TTGGTGTCTTACAGCCAGAGAGATCCGGACATGTTCCACTCGGCGCTCTTCGTGGACAC 480
      |||||
DB      141 LeuValPheTyrSerIleGlnIleArgAspMetPheHisLeuArgLeuPheValAspThr 160
QY      481 GACTCCAGCGTCAAGGTGTCTCGAAGAGTTCGCCGAGCGTGCAGGAGGAGGACCTG 540
      |||||
DB      161 AspSerAspValArgLeuSerArgValIleuArgAspValArgArgGIyArgAspLeu 180
QY      541 GAGCAGATTCTGACGCGAGTACACACACTTCGTGAACCGCGCTTCGAGGAGTTTGCCTG 600
      |||||
DB      181 GluGlnIleLeuThrGlnTyrThrThrPheValLysProAlaPheGIuGluPheCysLeu 200
QY      601 CCGACAAAGAAGATGCCGATGTGATCATCCACGAGAGGTGACAAATGTGTTGCCATC 660
      |||||
DB      201 ProThrLysArgTyrIleAspValIleIleProArgGIyValAspAsnMetValAlaIle 220
QY      661 AACCTGATCGTGACGACATCCAGACATTCGAAATGTGATCATGTGCAATGTGACACCA 720
      |||||
DB      221 AsnIleIleValGlnHisIleGlnAspIleLeuAsnGIyAspIleCysIyStrPheAsn 240
QY      721 GAGAGGTCCAAATGGCGGAGCTTACAAAGCGACCTTTTGTAGCCAGGAGCAACCTGGG 780
      |||||
DB      241 GLyGIySerAsnGIyArgSerTyrLysArgThrPheSerGIuProGIyAspHisProGIy 260
QY      781 ATCTGACCTCTGGGCAACGGTTCACATTTTGGAGTCCAGACGACGACCCAC 831
      |||||
DB      261 MetLeuThrSerGIyLysArgSerHisLeuGluSerSerArgProHis 277

RESULT 2
AAB93941
ID      AAB93941 standard; Protein; 277 AA.
XX
AC      AAB93941;
XX
DT      26-JUN-2001 (first entry)
XX
DE      Human protein sequence SEQ ID NO:13952.
XX
KW      Human; primer; detection; diagnosis; antisense therapy; gene therapy.
OS      Homo sapiens.
XX
PN      EP1074617-A2.
XX
PD      07-FEB-2001.
XX
PF      28-JUL-2000; 2000EP-0116126.
XX
PR      29-JUL-1999; 99JP-0248036.
PR      27-AUG-1999; 99JP-0300253.
PR      11-JAN-2000; 2000JP-0118776.
PR      02-MAY-2000; 2000JP-0183767.
PR      09-JUN-2000; 2000JP-0241899.
XX
PA      (HELI-) HELIX RES INST.
XX
PI      Ota T, Isogai T, Nishikawa T, Hayaashi K, Saito K, Yamamoto J;
PI      Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
```

DR WPI; 2001-318749/34.
 XX Primer sets for synthesizing polynucleotides, particularly the 5602
 PT full-length cDNAs defined in the specification, and for the detection
 PT and/or diagnosis of the abnormality of the proteins encoded by the
 PT full-length cDNAs -
 XX
 XX
 PS Claim 8; SEQ ID 13952; 2537bp + CD ROM; English.
 CC The present invention describes primer sets for synthesizing 5602
 CC full-length cDNAs defined in the specification. Where a primer set
 CC comprises: (a) an oligo-oligo primer and an oligonucleotide complementary
 CC to the complementary strand of a polynucleotide which comprises one of
 CC the 5602 nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in
 CC the specification. The primer sets can be used in antisense therapy and
 CC in gene therapy. The primer sets are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
 CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
 CC represent oligonucleotides, all of which are used in the exemplification
 CC of the present invention.
 CC
 SO Sequence 277 AA;

Alignment Scores:
 Pred. No.: 1.5e-112 Length: 277
 Score: 1450.00 Matches: 277
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 95.77% Indels: 0
 DB: 22 Gaps: 0

US-09-896-522-3 (1-834) x AAB93941 (1-277)

QY 1 ATGGCTTCGGCGGAGCCGGAAGCTCGCAGAGCCCGCGGAGCCGACCTCCGCAC 60
 DB 1 MetAlaSerAlaGlyGluAspCysGluSerProAlaProGluAlaAspArgProHis 20
 QY 61 CAGCGGCGCTTCCTGATAGGGGTGAGCGGCGGCGACCTGCGAGCGGAGAGTCGACCGGTGT 120
 DB 21 GluArgProPheLeuIleGlyValSerGlyGlyThrAlaSerGlyLeuSerThrValCys 40
 QY 121 GAGAGATCATGAGAGTTGCTGAGCAGAAAGAGTGAACAGCGGCGAGCGGAGAGTGTCTC 180
 DB 41 GluGlyIleMetGluLeuLeuGluGlnAlaGluGluAlaArgGlnArgGlyValVal 60
 QY 181 ATCTTACGCGAGCAGCGGTTCTTCAAGGTCCTGACGCGCAGAGCAGAAAGCGCTTGG 240
 DB 61 ILeuSerGlnAspArgPheTyrValLeuThrAlaGluGlnIleValAlaValAlaLeu 80
 QY 241 AAGAGCAGTACATTTGATGACCATCGAGATGCTTGTATGATGATTTGATGACAGAGCT 300
 DB 81 LysGlyGlnIleThrAsnIleAspHisProAspAlaAlaAspAsnAspLeuThrValArgThr 100
 QY 301 CTGAGAGACATCGTGGAGGCGAAACGCTGAGAGTCCGACCTATGATTTTGTGACACAC 360
 DB 101 LeuLysAsnIleValGluGlyLysThrValGluValProThrTyrAspPheValThrHis 120
 QY 361 TCAAGGTTACCGAGACCAACGCGGTGCTTACCTCGCGAGCTGGTTCTGTTTGAAGGCATC 420
 DB 121 SerArgLeuProGluThrThrValValTyrProAlaAspValValAlaLeuPheGluGlyIle 140
 QY 421 TTGTTGTTCTACAGCAGAGAGATCCGAGCATGTTCACACTGGCTCTTCTGTGACAC 480

DB 141 LeuValPheTyrSerGlnIleValLeuArgAspMetPheHisLeuArgLeuPheValAspThr 160
 QY 481 GACTCCGACGTACAGGCTGCTTCGAAAGATTCTCCGGAGCTGCCCGGAGGAGGACCTTG 540
 DB 161 AspSerAspValArgLeuSerArgValLeuArgAspValArgGlyLysArgAspLeu 180
 QY 541 GAGCAGATTCGACGACGATACCACTTCGGAAGCGCGCTTCGAGAGAGTTTGGCTTG 600
 DB 181 GluGlnIleLeuThrGlnIleThrThrPheValLysProAlaPheGluGluThrCysLeu 200
 QY 601 CCGACAAAGAGATGATCCGATGATGATCCACAGAGAGTGAACATATGTTGGTTCATC 660
 DB 201 ProThrLysLysTyrAlaAspValIleIleProArgGlyValAlaAspAsnMetValAlaIle 220
 QY 661 AACCTGATGTCGACGACATCCAGACATTTGAAATGTGATCATGTCAAATGACACGA 720
 DB 221 AsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysLysThrPHisArg 240
 QY 721 GAGAGGTCATATGGGCGGAGCTTACAAAGCGGAGCTTTTTCGACCCAGGAGACCTGGG 780
 DB 241 GlyGlySerAsnGlyArgSerTyrLysArgThrPheSerGluProGlyAspHisProGly 260
 QY 781 ATGCTGACCTCTGGCAACGCTCACATTTGAGTCCAGACGAGACCCAC 831
 DB 261 MetLeuThrSerGlyLysArgSerHisLeuGluSerSerSerArgProHis 277
 RESULT 3
 ID AAE16592 standard; Protein; 277 AA.
 AC AAE16592;
 XX
 XX 18-APR-2002 (first entry)
 DE Human 57658 protein.
 XX
 XX Human; uridine kinase-like protein; haematopoietic neoplastic disorder;
 KW 57658 protein; leukaemia; haematopoietic disorder; diabetes mellitus;
 KW rheumatoid arthritis; psoriasis; dermatitis; osteoporosis; rickets;
 KW sarcoma; myocardial infarction; hypertension; atherosclerosis;
 KW Alzheimer's disease; Parkinson's disease; Huntington's disease;
 KW gene therapy; chromosome mapping; tissue typing; dermatological;
 KW cytostatic; osteopathic; cardiac; neuroprotective; nootropic;
 KW anticovulant.
 KW
 XX Homo sapiens.
 XX
 FH Key
 FT Modified-site
 FT /label= N-myristoylation_site
 FT 25..151
 FT Domain
 FT /note= "Kinase uridine transferase ATP-binding
 FT phosphoribulokinase monophosphokinase precursor PRK
 FT cycle phosphopentokinase domain"
 FT 25..231
 FT Domain
 FT /note= "Phosphoribulokinase domain; Protein kinase
 FT family domain"
 FT 27..32
 FT Modified-site
 FT /label= N-myristoylation_site
 FT 30..37
 FT Binding-site
 FT /label= ATP/GTP-binding_site
 FT 34..36
 FT Modified-site
 FT /label= Protein_kinase-C_phosphorylation_site
 FT 38..41
 FT Modified-site
 FT /label= Casein_kinase_II_phosphorylation_site
 FT 100..102
 FT Modified-site
 FT /label= Protein_kinase-C_phosphorylation_site
 FT 108..115
 FT Modified-site
 FT /label= Tyrosine_kinase_phosphorylation_site
 FT 154..277
 FT Domain
 FT /note= "Kinase uridine monophosphokinase transferase
 FT ATP-binding kinase-like ribonucleoside pyrimidine FIS

FT Modified-site CDNA domain"

FT /label= Casein_kinase_II_phosphorylation_site 160..163

FT Modified-site /label= Casein_kinase_II_phosphorylation_site 167..169

FT Modified-site /label= Protein_kinase-C_phosphorylation_site 202..204

FT Modified-site /label= Protein_kinase-C_phosphorylation_site 241..246

FT Modified-site /label= N-myristoylation_site 247..249

FT Modified-site /label= Protein_kinase-C_phosphorylation_site 251..254

FT Modified-site /label= Casein_kinase_II_phosphorylation_site 260..265

FT Modified-site /label= N-myristoylation_site 264..266

FT Modified-site /label= Protein_kinase-C_phosphorylation_site 264..267

FT Modified-site /label= Amdation_site 268..271

FT Modified-site /label= Casein_kinase_II_phosphorylation_site 273..275

FT Modified-site /label= Protein_kinase-C_phosphorylation_site

PN W0200202761-A2.

XX 10-JAN-2002.

XX 28-JUN-2001; 2001WO-US21063.

XX 30-JUN-2000; 2000US-216503P.

XX (MILL-) MILENNIUM PHARM INC.

XX Glucksmann MA;

XX WPI; 2002-140091/18.

DR N-PSDB; AAD27186.

XX New isolated human uridine kinase family polypeptide 57658, useful for treating hematopoietic neoplastic disorders and disorders of neurons, heart and blood vessels

XX Claim 4; Fig 1a; 103pp; English.

XX The patent discloses human uridine kinase-like polypeptides, designated 57658 and polynucleotides encoding such proteins. 57658 DNAs and proteins are useful for developing diagnostic and therapeutic agents for 57658-mediated or related disorders such as hematopoietic neoplastic disorders (e.g. leukemias), hematopoietic disorders (e.g. psoriasis, dermatitis, diabetes mellitus, rheumatoid arthritis, disorders of bone (e.g. rickets, osteoporosis), heart (e.g. sarcoma, myocardial infarction), blood vessels (e.g. hypertension, atherosclerosis) and neurons (e.g. Huntington's disease, Alzheimer's disease, Parkinson's disease). 57658 sequences or their antibodies are useful in screening assays, detection assays (e.g. forensic biology) and predictive medicine (e.g. diagnostic assays, prognostic assays, and monitoring clinical trials and pharmacogenomics). They are useful as reagents for diagnosing and treating 57658-mediated disorders. 57658 DNAs are useful in gene therapy, in chromosome mapping, to identify an individual from a minute biological sample (tissue typing) and to aid in forensic identification of the biological sample. The present sequence is human 57658 protein.

XX Sequence 277 AA;

Alignment Scores:

Pred. No.: 1.5e-112 Length: 277

Score: 1450.00 Matches: 277

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 95..77% Indels: 0

DB: 23 Gaps: 0

US-09-896-522-3 (1-834) x AAB16592 (1-277)

QY 1 ATGGCTTCGGCGGAGCGGAGACTGCGAGAGCCCCCGGAGGCCACCGTCCGCAC 60

Db 1 MetAlaSerAlaGlyGlyGlnAspCysGlnSerProAlaProGlnAlaAspArgProHis 20

QY CAGCGGCGCTTCCTGATGAGGGGTGAGCGGCGGACACTGCGAGGGGAGTCAGCCGTGT 120

Db 21 GlnArgProPheLeuIleGlyAlaSerGlyGlyThrAlaSerGlyLysSerThrValCys 40

QY 121 GAGAGATCATGAGAGTTGCTGGAGAGAAAGAGTGAACACCGGACGGAGGAGTGTGTC 180

Db 41 GlnLysIleMetGlnLeuLeuGlnGlnAlaGlnGlnAlaGlnGlnAlaArgValVal 60

QY 181 ATCTTACGACGAGACAGGTTTCTACAGGTCCTTGAACGCGGACGAGAGGCGCTTG 240

Db 61 IleLeuSerGlnAspArgPheThrLysValLeuThrAlaGlnGlnLysAlaLysAlaLeu 80

QY 241 AAGAGACATGACATTTTACCATCCAGATGCGCTTGAATGATTTGATGACAGACT 300

Db 81 LysGlyGlnThrAsnPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThr 100

QY 301 CTGAGAGACATCGTGGAGGCGAAAGCGTGAAGTGCAGCTATGATTTTGACACAC 360

Db 101 LeuLysAsnIleValGlnGlyLysThrValGlnValProThrLysPheValThrHis 120

QY 361 TCAGGTTACAGAGACACAGGTGTCTACCTCGCGGACGTGGTCTGTTGAGGCGATC 420

Db 121 SerArgLeuProGlnThrValValTyrProAlaAspValValLeuPheGlnGlyLe 140

QY 421 TTGCTGTTCTACAGCGGAGAGATCCGAGACATGTTCCACTCGGCTCTCGTGACACC 480

Db 141 LeuValPheThrLysSerGlnGlnIleArgAspMetPheHisLeuArgLeuPheValAspThr 160

QY 481 GACTCGACGTCAGGCTGTCTCGAAGAGTTCTCCGGACGTGGCGCCGAGGAGGACCTG 540

Db 161 AspSerAspValArgLeuSerArgValLeuArgAspValArgGlyLysArgLeu 180

QY 541 GAGCAATTCGACGCGAGTACCACTTCGTGAGACCGGCGCTCGAGGAGTTCTGCCCTG 600

Db 181 GlnGlnIleLeuThrLysThrThrPheValLysProAlaPheGlnIleLysPheCysLeu 200

QY 601 CCGACAAAGAGATGCCGATGATCATCCACGAGAGTGGACAAATGTTGTCATC 660

Db 201 ProThrLysLysThrAlaAspValIleIleProArgGlyValAspAsnMetValAlaIle 220

QY 661 AACCTGATCGTGCAGACATCCAGACATTCGATGTGATGTGACATTCGAATGGACCGA 720

Db 221 AsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysLysTrpHisArg 240

QY 721 GAGGGTCCAAATGGGCGGAGCTTCAAGCGGACCTTTCTGAGCCAGAGGAGACCACTGGG 780

Db 241 GlyGlySerAsnGlyArgSerThrLysArgThrPheSerGlnProGlyLysAspProGly 260

QY 781 ATCTGACCTTCGCAACGCTCATCTTGGAGTCCAGACGAGACCCAC 831

Db 261 MetLeuThrSerGlyLysArgSerHisLeuGlnSerSerThrProHis 277

RESULT 4

ID AAM41288

AC AAM41288 standard; Protein; 296 AA.

XX AAM41288;

DT 22-OCT-2001 (first entry)

XX Human polypeptide SEQ ID NO 6219.

DE Human; noctropic; immunosuppressant; cytostatic; gene therapy; cancer; peripheral nervous system; neuropathic; central nervous system; CNS; Alzheimer's; Parkinson's disease; Huntington's disease; hemostatic; KM amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemoclastic; KM chemokinetic; thrombolytic; drug screening; arthritis; inflammation;

KX	leukaemia.
XX	
OS	Homo sapiens.
XX	
PV	MO200153312-A1.
XX	
PD	26-JUL-2001.
XX	
PF	26-DEC-2000; 2000WO-US34263.
XX	
PR	21-JAN-2000; 2000US-0488725.
PR	25-APR-2000; 2000US-0552317.
PR	09-JUL-2000; 2000US-0598042.
PR	19-JUL-2000; 2000US-0620312.
PR	03-AUG-2000; 2000US-0653450.
PR	14-SEP-2000; 2000US-0662191.
PR	19-OCT-2000; 2000US-0693036.
XX	29-NOV-2000; 2000US-0727344.
PA	(HYSE-) HYSEQ INC.
XX	
P1	Tang Y T, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
P1	Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
P1	Zhao QA, Zhou P, Goodrich R, Drmanec RT;
XX	
DR	WPI; 2001.442253/47.
XX	N-PDB; AAI60444.
PT	
PT	Novel nucleic acids and polypeptides, useful for treating disorders
XX	such as central nervous system injuries -
PS	
XX	Example 2; SEQ ID NO 6219; 10078pp; English.
CC	The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC	the encoded polypeptides (AAM38642-AAM42213) with neurotropic,
CC	immunosuppressant and cyostatic activity. The polynucleotides are useful
CC	in gene therapy. A composition containing a polypeptide or polynucleotide
CC	of the invention may be used to treat diseases of the peripheral nervous
CC	system, such as peripheral nervous injuries, peripheral neuropathy and
CC	localized neuropathies and central nervous system diseases, such as
CC	Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC	lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC	utilization of the activities such as: Immune system suppression,
CC	Activin/inhibin activity, chemoclastic/chemokinetic activity, haemostatic
CC	and thrombotic activity, cancer diagnosis and therapy, drug screening,
CC	assays for receptor activity, arthritis and inflammation, leukaemias and
CC	C.N.S disorders.
CC	Note: The sequence data for this patent did not form part of the printed
CC	specification.
XX	
SQ	Sequence 296 AA;
Alignment Scores:	
Pred. No.:	1.53e-112 Length: 236
Score:	1450.00 Matches: 277
Percent Similarity:	100.00% Conservative: 0
Best Local Similarity:	100.00% Mismatches: 0
Query Match:	95.77% Indels: 0
DB:	22 Gaps: 0
US-09-896-522-3 (1-834) x AAM41288 (1-296)	
OY	1 ATGGCTTCGCGGAGCGGAAGACTCCAGAGCCCGCGCGGAGGCGACCGTCCGAC 60
Dd	
20 MetcLaserlaaglygyluaepCySglusePrOalPProGlunlaasprgrProtis 39	
OY	61 CAGCGGCCCTTCGTATGAGGGGTGAGCGCGGCACCTCCAGCGGGAAGTGACCGTGTGT 120
Dd	
40 GlnarGProheLeuIlleglyValSerGlyGlyThralaSerGlyLyseerThralValCys 59	
OY	121 GAGAAGTCAATGAGTTCCTGGACAAGAACGAGGTGGAACACCGCAGCGGAAGTGTCT 180
Dd	
60 GlunylelleetGluLeuLeuenglyGlnasGluValaGluGlnarGGLnarGlyValVal 79	

QY	181	ATCCGACCCAGGACAGGCTTCTACAGGCTCCGACGGGACAGAGCCAGAGGCTTG	240
Db	80	lleuSerSerGlnAspAlaPheTyrIysValIleuThrAlaGlnGlnIysAlaIysAlaIeu	99
QY	241	AAAGGACAGTCAATTTTGGACCATCCAGATGCGCTTGTATATGATTTGATGACAGAGCT	300
Db	100	LysGlyGlnTyrAenPheAspHisProAspAlaPheAspAsnAspIleuMetHisArgThr	119
QY	301	CTGAAGAACATCGGTGGAGGGCAAAACGGTGGAGGGTCCGACCTATGATTTTGTGACACAC	360
Db	120	LeuIysAsenIleValGlnGlyIysThrValGluValProThrTyrAspPheValThrHis	139
QY	361	TCAAGGTTACACAGAGACACCGGTGGTCTAACCTCGGAGACGTGGTCTGTGTTGAGGGCATC	420
Db	140	SerAlaGluLeuProGlnThrThrValValIlyrProAlaAspValValIleuPheGlnGlyIle	159
QY	421	TTGTGTGTTCTACAGCCAGAGAGATCCGGAGACATGTTCCACTGCGCCTTCTGTGACACCC	480
Db	160	LeuValPheTyrSerGlnGlnIleArgAspMetPheHisIleuArgIleuPheValAspThr	179
QY	481	GACTCCGACGTCAGGCTGTCTTGAAAGATTTCTCCGGAGACGTGCCCGAGAGAGGACCTG	540
Db	180	AspSerAspValAlaGluLeuSerArgThrValIleuAlaGAspValAlaArgGlyAlaGAspLeu	199
QY	541	GAGCAGATTCCTGACGACGAGTACACACACTTCGTAAGACCGGACCTTGACAGAGATTCGCTG	600
Db	200	GlnGlnIleIleuThrGlnTyrThrThrPheValIysProAlaPheGlnIuPheCysIeu	219
QY	601	CCGACAAAGAAATATCCGATGTGATCATCCACGAGAGAGTGAACAAATATGTTGCCATC	660
Db	220	ProThrIlyslsIyAlaIAspValIleIleProArgGlyValIAspAsnMetValAlaIle	239
QY	661	AACCTGATCTGTGCAGACATATCCAGACATTCCTGAATGTGTGAATCTTGCAATATGGACCCGA	720
Db	240	AsnIleuIleValIleGlnHisIleGlnAspIleLeuAsnGlyAspIleCysIysThrPHisArg	259
QY	721	GGAGGGTCCATATGGGCGAGGCTACAGCGGAGCACTTTTCTGAGCCAGGGGACACCTGGG	780
Db	260	GlyIylsSerAsnGlyIysArgSerTyrIysArgThrPheSerGluProGlyAspHisProGly	279
QY	781	ATGCTGACCTCTGGCAAAACGGTCACATTTGGAGTCCAGACAGACACCCAC	831
Db	280	MetIeuThrSerGlyIysArgSerHisIleuGlnSerSerArgProHis	296
RESULT 5			
AAAG64506			
ID	AAAG64506	standard; Protein; 276 AA.	
AC	AAAG64506;		
XX	02-OCT-2001	(first entry)	
XX	Human uridine kinase.		
XX	Human; uridine kinase; UK.		
OS	Homo sapiens.		
XX	CNI287172-A.		
XX	14-MAR-2001.		
XX	07-SEP-1999;	99CN-0118818.	
XX	07-SEP-1999;	99CN-0118818.	
XX	07-SEP-1999;	99CN-0118818.	
XX	(UYFU-) UNIV FUDAN.		
XX	Yu L, Zhao Y, Zhang H;		
XX	WPI; 2001-409529/44.		
XX	N-PSDB; AAH75355.		

XX Human uridine kinase and its coding sequence, preparation and
PT application -
PS Claim 2, Page 15-16(Disclosure); 20pp, Chinese.
XX The invention relates to human uridine kinase (UK).
CC
XX
SQ Sequence 276 AA;

Alignment Scores:
Pred. No.: 4.39e-106 Length: 276
Score: 1372.50 Matches: 266
Percent Similarity: 97.11% Conservative: 3
Best Local Similarity: 96.03% Mismatches: 7
Query Match: 90.65% Indels: 1
DB: 22 Gaps: 1

US-09-896-522-3 (1-834) x AA664506 (1-276)

QY 1 ATGCTTCGCGGAGGAGGAGAGCTGCGAGCCCGCGGAGCCGACCTCCGAC 60
D 1 MetAlaSerAlaGlyGlyGlyGlyAlaArgAlaArgAla---GlyAlaAsnArgProHis 19
QY 61 CAGCGCCCTTCCTGATAGGGGTGAGCGCGCACTGCCAGCGGAACTCCAGCTGTGT 120
D 20 GlnArgProPheLeuIleGlyValSerGlyThrAlaSerGlyLysSerThrValCys 39
QY 121 GAGAGATCATGAGTTCCTGGACAGACAGAGGTGGAACGCGGCGGAGAGTGTTC 180
D 40 GluLysIleMetGluLeuGluGlnAsnGluValGluGlnArgHisGlyLysValVal 59
QY 181 ATCTGAGCCGAGCAGAGTTCACAAAGGTCTGACGCGAGAGCAGAAAGCCGACCTTG 240
D 60 IleLeuSerGlnAspArgPheThrLysValLeuThrAlaGluGlnLysAlaLysAlaLeu 79
QY 241 AAAGAGACATCAATTTTGCACATCCAGATGCTTTGATTAATGATTGATGCAGAGACT 300
D 80 LysGlyGlnTyrAsnPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThr 99
QY 301 CTGAGAGAACATCGTGGAGGGGAAAACGGTGGAGGTGCCGACCTTGTATTTGTACACAC 360
D 100 LeuLysAsnIleValGluGlyLysThrValGluValProThrTyrAspPheValThrHis 119
QY 361 TCAAGTTACCGAGACACACGCTGCTCACTCCGCGACGTGTTCTGTTGAGGGCATC 420
D 120 SerArgLeuProGluThrThrValValTyrProAlaAspValValLeuPheGluGlyIle 139
QY 421 TTGGTGTCTACAGCCAGAGATCCGGACATGTTCACCTGCGCTCTTCTGGAGACAC 480
D 140 LeuValPheTyrSerGlnGluIleArgAspMetPheHisIleuArgLeuPheValAspThr 159
QY 481 GACTCCGACGTCAGGCTGCTCCGAAAGATTCTCCGGACGTCGCGCCGAGGGAGGACTTG 540
D 160 AspSerAspValAlaGluSerArgArgValLeuArgAspValAlaArgGlyAlaArgAspLeu 179
QY 541 GAGCAGATTTCAGCAGCATACACACCTTCGTAAGCGCGGCTTCGAGAGATTTCGCTTG 600
D 180 GluIleIleLeuThrGlnTyrThrThrPheValLysProAlaPheGluGluPheCysLeu 199
QY 601 CCGACAAAGAGATGCCGATGTATCATCCACGAGAGGTGACCAATATGTTGCCATC 660
D 200 ProThrLysLysTyrAlaAspValIleIleProArgGlyValAlaAsnMetValAlaIle 219
QY 661 AACCTGATCGCGACATCCAGACATTTGATAGTGTGACATCTGCAATGGCACCGA 720
D 220 AsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysLysTyrPheArg 239
QY 721 GAGAGGTCGAATGGGCGGAGTACAAAGGAGCATTTTCTGAGCGCAGGGGACACCTGGG 780
D 240 GlyLysSerAsnGlyArgThrTyrLysArgThrPheSerGlnProGlyAspHisProGly 259
QY 781 ATGTGACCTCTGGCAACGGTCAATTGGAGTCCAGACAGACAGACCCAC 831

DB 260 MetLeuThrSerGlyLysArgSerHisLeuGluSerSerArgProHis 276

RESULT 6
AA014412
ID AA014412 standard; Protein; 260 AA.
XX
XX AA014412;
AC
XX
DT 02-MAY-2002 (first entry)
XX
XX
DE Protein of a human uridine kinase (UDK).
XX
XX Human uridine kinase; diagnostic assay; mutation detection; UDK;
KW probe; chromosome localisation study; tissue expression; gene therapy;
KW antibody; vaccine; human ovarian cancer; immunological disorder;
KW human colon carcinoma; immunogen.
XX
XX Homo sapiens.
OS
XX
XX W0200172963-A2.
XX
XX PD 04-OCT-2001.
XX
XX PF 27-MAR-2001; 2001WO-US09663.
XX
XX PR 27-MAR-2000; 2000US-0536647.
XX
XX (SMIRK) SMITHKLINE BEECHAM CORP.
XX
XX PI Ho YS, Johnson RK;
XX
XX WPI: 2001-626259/72.
DR N-PSDB; AAK98735.
XX
XX
PT Novel human uridine kinase polypeptides useful for treating cancers,
PT and to identify agonists and antagonists of the polypeptide useful for
PT treating conditions associated with uridine kinase imbalance -
XX
PS Claim 3; Page 29-30; 31pp; English.
XX
XX The invention relates to newly identified human uridine kinase (UDK)
CC polypeptides and polynucleotides and methods for producing such
CC polypeptides by recombinant techniques. Also disclosed in the invention
CC are methods for utilising uridine kinase polypeptides and polynucleotides
CC in diagnostic assays. The polynucleotides and polypeptides of the
CC invention may be used as diagnostic reagents by detecting mutations in an
CC associated gene. An array of oligonucleotide probes comprising the
CC uridine kinase polynucleotide sequence or fragments thereof can be
CC constructed to conduct efficient screening of genetic mutations, for
CC example. Detection of abnormally decreased or increased levels of
CC polypeptide or mRNA expression may also be used for diagnosing or
CC determining susceptibility of a subject to a disease of the invention.
CC The polynucleotide sequences of the invention can be used for chromosome
CC localisation studies and tissue expression studies. The polypeptides of
CC the invention or fragments thereof may be used as immunogens to produce
CC antibodies. These antibodies may be employed to isolate or identify
CC clones expressing the polypeptide. The polypeptides and polynucleotides
CC of the invention can be used as a vaccine or in gene therapy to treat
CC diseases such as human ovarian cancer, human colon carcinomas, and
CC immunological disorders. This sequence represents the protein of a human
CC uridine kinase of the invention.
XX
SQ Sequence 260 AA;

Alignment Scores:
Pred. No.: 2.67e-105 Length: 260
Score: 1363.00 Matches: 260
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 90.03% Indels: 0
DB: 22 Gaps: 0

US-09-896-522-3 (1-834) x AA014412 (1-260)

```

QY 1 ATGGCTTCGGCGGGAAGCGGAGACCGCCGCGGAGCGGACCGTCCGAC 60
Db 1 MetAlaSerIaGlyGluAspCyGluSerProAlaProGluAlaAspArgProHis 20
QY 61 CAGCGGCCCTTCCTGATAGGGGTGAGCGGCGGACCTCCAGCGGGAAGTGCACCGTGTGT 120
Db 21 GlnArgProPheLeuIleGlyValSerGlyGlyThrAlaSerGlyLysSerThrValCys 40
QY 121 GAGAGATCATGAGATTGCTGGGACAGAACGAGGTGGAACGCGGACCGGAAGTGTCTC 180
Db 41 GlnLysIleMetGluLeuLeuGlnAsnGlnValGluGlnArgGlnArgLysValVal 60
QY 181 ATCTGAGCGGACGAGCTTCAAGGTCCTGACGCGGACAGGACGGAAGCGCTTGG 240
Db 61 IleuSerGlnAspArgPheTyrLysValLeuThrAlaGluGlnLysAlaLysAlaLeu 80
QY 241 AAAGCAGTACATTTTGAACCATCCAGATGCTTTGATATGATTGATGACAGGACT 300
Db 81 LysGlyGlnIleTyrAsnProPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThr 100
QY 301 CTGAGAACATCGTGGAGGCGCAAAACGCTGAGAGTGCCTGACTTATGATTTTGTGACACAC 360
Db 101 LeuLysAsnIleValGluGlyLysThrValGluValProThrTyrAspPheValThrHis 120
QY 361 TCAAGGTTACCAAGACACGAGTGTCTACCTCCGACGCTGTCTCTTGTGAGGGATC 420
Db 121 SerArgLeuProGluThrValThrValValTyrProAlaAspValValLeuPheGluGlyLe 140
QY 421 TTGGTGTTTACAGCCGAGGATCCGGGACATGTTTCCAGCTGGCGCTCTTCTGAGGACAC 480
Db 141 LeuValPheTyrSerGlnGlnIleArgAspMetPheHisLeuArgLeuPheValAspThr 160
QY 481 GACTCCGACGTCAAGCTGTCTCGAAGATTCTCCGGGACGTCGCCGAGGAGGAGACTG 540
Db 161 AspSerAspValArgLeuSerAspArgValLeuArgAspValArgArgGlyArgAspLeu 180
QY 541 GACCAAGTTTTCAGCGCATCACCACTTGTGTAAAGCGCGCTTTCGAGGAGTCTGCTG 600
Db 181 GlnGlnIleLeuThrGlnTyrThrThrPheValLysProAlaPheGluGlnPheCysLeu 200
QY 601 CCGACAAGAAGTATGCGGATGATCATCCACGAGAGAGTGGACATATGTTGTCATC 660
Db 201 ProThrLysLysTyrAlaAspValIleIleProArgGlyValAspMetValAlaIle 220
QY 661 AACCTGATCGTGACGACATCCAGGACATTTCTGAATGTGACATCTGCAATGGCACCGA 720
Db 221 AsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlnLysPheCysLysThrPheArg 240
QY 721 GAGAGGTCCAAATGGCGGAGCTACAGCGGACCTTTTCTGAGCCAGGGGACACACCTGGG 780
Db 241 GlyGlySerAsnGlyArgSerTyrLysArgThrPheSerGluProGlyAspHisProGly 260

```

RESULT 7

ID AA014413 standard; Protein; 277 AA.

AC AA014413;

DT 02-MAY-2002 (first entry)

DE Protein relating to a human uridine kinase (UDK) of the invention.

XX Human; uridine kinase; diagnostic assay; mutation detection; UDK;

KM probe; chromosome localisation study; tissue expression; gene therapy;

KM antibody; vaccine; human ovarian cancer; immunological disorder;

XX human colon carcinoma; immunogen.

OS Unidentified.

XX WO200172963-A2.

```

PD 04-OCT-2001.
XX
PF 27-MAR-2001; 2001MO-US09663.
XX
XX 27-MAR-2000; 2000US-0536647.
PR
XX
PA (SMIK ) SMITHKLINE BEECHAM CORP.
PI Ho YS, Johnson RK,
XX
XX WPI, 2001-626259/72.
DR
XX
XX Novel human uridine kinase polypeptides useful for treating cancers,
PT and to identify agonists and antagonists of the polypeptide useful for
PT treating conditions associated with uridine kinase imbalance
XX
XX Disclosure; Page 23; 31pp; English.
XX
XX The invention relates to newly identified human uridine kinase (UDK)
CC polypeptides and polynucleotides and methods for producing such
CC polypeptides by recombinant techniques. Also disclosed in the invention
CC are methods for utilising uridine kinase polypeptides and polynucleotides
CC in diagnostic assays. The polynucleotides and polypeptides of the
CC invention may be used as diagnostic reagents by detecting mutations in an
CC associated gene. An array of oligonucleotide probes comprising the
CC uridine kinase polynucleotide sequence or fragments thereof can be
CC constructed to conduct efficient screening of genetic mutations, for
CC example. Detection of abnormally decreased or increased levels of
CC polypeptide or mRNA expression may also be used for diagnosing or
CC determining susceptibility of a subject to a disease of the invention.
CC The polynucleotide sequences of the invention can be used for chromosome
CC localisation studies and tissue expression studies. The polypeptides of
CC the invention or fragments thereof may be used as immunogens to produce
CC antibodies. These antibodies may be employed to isolate or identify
CC clones expressing the polypeptide. The polypeptides and polynucleotides
CC of the invention can be used as a vaccine or in gene therapy to treat
CC diseases such as human ovarian cancer, human colon carcinomas, and
CC immunological disorders. This sequence represents the protein relating to
CC a human uridine kinase (UDK) of the invention.
CC NOTE: The present sequence is elated as being the same as that shown as
CC SEQ ID NO: 2 in the sequence listing of the specification. However, the
CC sequences differ.
XX
XX
SQ Sequence 277 AA;

```

Alignment Scores:

Pred. No.: 4.04e-102 Length: 277
 Score: 1325.00 Matches: 254
 Percent Similarity: 95.67% Conservative: 11
 Best Local Similarity: 91.70% Mismatches: 12
 Query Match: 87.52% Indels: 0
 DB: 22 Gaps: 0

US-09-896-522-3 (1-834) x AA014413 (1-277)

```

QY 1 ATGGCTTCGGCGGGAAGCGGAGACCGCCGCGGAGCGGACCGTCCGAC 60
Db 1 MetAlaSerIaGlyGlyGlySerGluSerAlaIaProGluAlaAspArgProGln 20
QY 61 CAGCGGCCCTTCCTGATAGGGGTGAGCGGCGGACCTCCAGCGGGAAGTGCACCGTGTGT 120
Db 21 ProArgProPheLeuIleGlyValSerGlyGlyThrAlaSerGlyLysSerThrValCys 40
QY 121 GAGAGATCATGAGATTGCTGGGACAGAACGAGGTGGAACGCGGACCGGAAGTGTCTC 180
Db 41 GlnLysIleMetGluLeuLeuGlnAsnGlnValAspArgGlnArgLysValVal 60
QY 181 ATCTGAGCGGACGAGCTTCAAGGTCCTGACGCGGACAGGACGGAAGCGCTTGG 240
Db 61 IleuSerGlnAspArgPheTyrLysValLeuThrAlaGluGlnLysAlaLysAlaLeu 80
QY 241 AAAGCAGTACATTTTGAACCATCCAGATGCTTTGATATGATTGATGACAGGACT 300

```

Db LysGlyGlnItyrAsnPhenAspHisProAspAlaPheAspAsnLeuMetHisIleuThr 100
 Qy 301 CTGAAGAACATCGTGGAGGGGAAACGGTGGAGTCCGACCTTATGATTTTGTACACAC 360
 Db 101 LeuysAsnIleValIGluGlyLysThrValIGluValProThrItyrAspPheValThrHis 120
 Qy 361 TCAAGGTTACAGAGACACCGGTGTCTACCTCCGACGCTGTTCTGTTGAGGACATC 420
 Db 121 SerArgLeuProGluThrItyrThrValItyrProAlaAspValIleLeuPheGluGlyIle 140
 Qy 421 TTGGTGTTCACAGCCAGAGATCCGGACATGTTCCACCTGGCCCTCTTCTGTGACACC 480
 Db 141 LeuValPheItyrThrGlnGluIleAlaArgAspMetPheHisIleuArgLeuPheValAspThr 160
 Qy 481 GACTCCGAGCTGAGGCTGTCTCGAAGAGTTCTCCGGACGCGCCCGACGAGGAGGACCTG 540
 Db 161 AspSerAspValAlaArgLeuSerArgArgValIleuArgAspValIGlnArgGlyArgAspLeu 180
 Qy 541 GAGCAGATTCTGACGACATACACACCTTCGTTGAAAGCCGCGCTTCGAGAGTTCTGCCTG 600
 Db 181 GluGlnIleLeuThrGlnItyrThrAlaPheValItyrProAlaPheGluGluPheCysLeu 200
 Qy 601 CCGACAAAGAAATGTCGGATGTATCATCCACGAGAGTGAACAATATGTTGCCATC 660
 Db 201 ProThrItyrLysItyrAlaAspValIleIleProArgGlyValAspAsnMetValAlaIle 220
 Qy 661 AACCTGATCTGCGACATCCAGACATCTGATGATGTTGATCGAATATGGACCGA 720
 Db 221 AsnIleuIleValIGlnHisIleGlnHisPheIleuAsnGluLysAspLeuCysLysArgHisArg 240
 Qy 721 GGAGGGTCCAAATGGCGGAGCTACAGGCGACCTTTCTGAGCGAGGACACCTCTGG 780
 Db 241 GlyGlyProAsnGluArgAsnHisIleuArgThrPheProGluProGlyAspHisProGly 260
 Qy 781 ATGCTGACCTTGGCAACCGTCACTTTGAGTTCAGACAGACACCCAC 831
 Db 261 ValIleuAlaThrGlyLysArgSerHisIleuGlnIleuSerSerArgProHis 277
 RESULT 8
 ABB89353
 ID ABB89353 standard; Protein; 190 AA.
 AC ABB89353;
 XX
 DT 24-MAY-2002 (first entry)
 XX
 DE Human polypeptide SEQ ID NO 1729.
 XX
 KM Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
 KM antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor;
 KM valnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
 KM cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KM neurological disease; infection; human; secreted protein.
 XX
 OS Homo sapiens.
 XX
 MO20010304-A2.
 XX
 PD 29-NOV-2001.
 XX
 PF 18-MAY-2001; 2001WO-US16450.
 XX
 PR 19-MAY-2000; 2000US-205515P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Birse CE, Rosen CA;
 DR WPI; 2002-122018/16.
 DR N-PSDB; ABL89762.
 XX
 PT Novel 1405 isolated polypeptides, useful for diagnosis, treatment and
 prevention of neural, immune system, muscular, reproductive,

PT gastrointestinal, pulmonary, cardiovascular, renal and proliferative
 PT disorders -
 XX
 PS Claim 11; SEQ ID NO 1729; 2081pp + Sequence Listing; English.
 XX
 CC The invention relates to novel genes (ABL9449-ABL90853) and proteins
 CC (ABB89040-ABB90444) useful for preventing, treating or ameliorating
 CC medical conditions e.g. by protein or gene therapy. The genes are
 CC isolated from a range of human tissues disclosed in the specification.
 CC The nucleic acids, proteins, antibodies and (ant)agonists are useful
 CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
 CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
 CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
 CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
 CC hemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
 CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
 CC colitis; (c) cardiovascular disorders such as myocardial ischemias;
 CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
 CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
 CC and parasitic infections.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at fcp.wipo.int/pub/published_pct_sequences.
 CC
 XX
 SQ Sequence 190 AA;
 Alignment Scores:
 Pred. No.: 5.73e-71 Length: 190
 Score: 951.00 Matches: 187
 Percent Similarity: 97.91% Conservative: 0
 Best Local Similarity: 97.91% Mismatches: 3
 Query Match: 62.81% Indels: 1
 DB: Gaps: 0
 US-09-896-522-3 (1-834) x ABB89353 (1-190).
 Qy 1 ATGCTTCCGCGGAGGCGAAGATCGAGACCCCGCGGAGCGACCTCCGAC 60
 Db 1 MetAlaSerIleGlyGluAspCysGluSerProAlaProIleuAlaAspArgProHis 20
 Qy 61 CACGGCCCTTCTGATPAGGGGTGAGCGGCGGACCTGCCAGCGGAAATGACCGTGT 120
 Db 21 GlnArgProPheLeuIleGlyValSerGlyItyr**SerGlySerThrValCys 40
 Qy 121 GAGAAGATCATGGAGTTCGCTGGGACAGAACAGAGCTGGAACAGCGGAGGAAAGTGGTC 180
 Db 41 GluLysIleMetGluLeuLeuGluGlnAsnGluValIGluGlnArgIleArgLysVal 60
 Qy 181 ATCTGAGCCAGAGACAGTCTTCAAGATCTTGACCGGACAGAGCAAGGCAAGGCTTG 240
 Db 61 IleuSerGlnAspArgPheItyrLysValIleuThrAlaGluGlnLysAlaLysAlaLeu 80
 Qy 241 AAGGACAGTACATTTTGAACATCCAGATGCTTTGATGATGATTTGATGCACAGACT 300
 Db 81 LysGlyGlnItyrAsnPhenAspHisProAspAlaPheAspAsnLeuMetHisIleuThr 100
 Qy 301 CTGAAGAACATCGTGGAGGGGAAACGGTGGAGTCCGACCTTATGATTTTGTACACAC 360
 Db 101 LeuysAsnIleValIGluGlyLysThrValIGluValProThrItyrAspPheValThrHis 120
 Qy 361 TCAAGGTTACAGAGACACCGGTGTCTACCTCCGACGCTGTTCTGTTGAGGACATC 420
 Db 121 SerArgLeuProGluThrItyrThrValItyrProAlaAspValIleLeuPheGluGlyIle 140
 Qy 421 TTGGTGTTCACAGCCAGAGATCCGGACATGTTCCACCTGGCCCTCTTCTGTGACACC 480
 Db 141 LeuValPheItyrThrGlnGluIleAlaArgAspMetPheHisIleuArgLeuPheValAspThr 160
 Qy 481 GACTCCGAGCTGAGGCTGTCTCGAAGAGTTCTCCGGACGCGCCCGACGAGGAGGACCTG 540
 Db 161 AspSerAspValAlaArgLeuSerArgArgValIleuArgAspValIGlnArgGlyArgAspLeu 180
 Qy 541 GACGAGATTCTGACGACATACACACCTTCGTTGAAAGCCGCGCTTCGAGAGTTCTGCCTG 571

Db 181 GluGlnIleuThrHis-ThrProProSer 190
 RESULT 9
 ID AAM38694 standard; Protein; 261 AA.
 XX AAM38694;
 AC
 XX
 DT 22-OCT-2001 (first entry)
 XX
 DE Human polypeptide SEQ ID NO 1839.
 XX
 KM Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KM peripheral nervous system; neuropathy; central nervous system; CNS;
 KM Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KM amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KM chemokine; thrombolytic; drug screening; arthritis; inflammation;
 KM leukaemia.
 XX
 OS Homo sapiens.
 PN WO20015312-A1.
 PD 26-JUL-2001.
 XX
 PF 26-DEC-2000; 2000MO-US34263.
 XX
 PR 21-JAN-2000; 2000US-0488725.
 PR 25-APR-2000; 2000US-0552317.
 PR 09-JUL-2000; 2000US-0598042.
 PR 19-JUL-2000; 2000US-0620312.
 PR 03-AUG-2000; 2000US-0653450.
 PR 14-SEP-2000; 2000US-0662191.
 PR 19-OCT-2000; 2000US-0693036.
 PR 29-NOV-2000; 2000US-0727344.
 XX
 PA (HYSE-) HYSEQ INC.
 PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Wehtman T, Xu C, Xue AJ, Yang Y, Zhang J;
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
 XX
 DR MPI; 2001-442253/47.
 DR N-PSDB; AAI57850.
 PT Novel nucleic acids and polypeptides, useful for treating disorders
 PT such as central nervous system injuries -
 XX
 PS Example 3; SEQ ID NO 1839; 10078bp; English.
 XX
 CC The invention relates to human nucleic acids (AAI57798-AAI61369) and
 CC the encoded polypeptides (AAM38642-AA42213) with nootropic,
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localized neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression.
 CC Activation/inhibition of chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemia and
 CC C.N.S disorders.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification.
 CC
 SQ Sequence 261 AA;

Alignment Scores: 1.66e-69 Length: 261
 Pred. No.: 934.00 Matches: 189
 Score:

Percent Similarity: 79.41% Conservative: 27
 Best Local Similarity: 69.49% Mismatches: 40
 Query Match: 61.69% Indels: 16
 DB: 22 Gaps: 4
 US-09-896-522-3 (1-834) x AAM38694 (1-261)
 QY 22 GACTCGAGAGAGCCCGCGGAGCCGACCGTCCGAC---CAGCGCCCTTCTGATTA 78
 DB 4 AepSerGlnGlnThrLeuGlnIleuHisGlnGlnProAsnGlyGlyGluProPheLeuIle 23
 QY GGGGTGAGCGCGGCGGACCTCCAGCGGAAAGTGCACCGTGTGAGAAGATCATGAGATTG 138
 DB 24 GlyValSerGlyGlyThrIleSerGlyLysSerSerValCysAlaLysIleValGlnLeu 43
 QY 139 CTGGGACAGAGAGAGGTGAGACAGCGGCGGAGAGGTGTCTATCTTGACCGACAGAG 198
 DB 44 LeuGlyGlnAsnGlnValAspTyrArgGlnLysGlnValValIleLeuSerGlnAspSer 63
 QY 199 TTCTACAAAGTCTCGACGCGAGAGGCGAAGGCGCTTGAAAGAGACATATTTT 258
 DB 64 PheTyrArgValLeuThrSerGlnLysAlaLysAlaLeuLysGlyGlnPheAsnPhe 83
 QY 259 GACCATCCAGATGCTTTGATATGATTTGATGACACGAGCTCTGAGAACAATCGTGAG 318
 DB 84 AspHisProAspAlaPheAspAsnGlnLeuIleLeuLysThrLeuLysGlnIleThrGln 103
 QY 319 GGCAGAAACGTTGAGGTCCGACCTATGATTTTGTGACACACTCAAGATTACAGAGACC 378
 DB 104 GlyLysThrValGlnIleProValTyrAspPheValSerHisSerArgLysGlnGluThr 123
 QY 379 ACGGTGCTTACCGTCCGAGCGGTCTGTTGAGGGCATCTGTGTTTCAACACCGAG 438
 DB 124 ValThrValTyrProAlaAspValValLeuPheGlnGlyIleLeuAlaPheTyrSerGln 143
 QY 439 GAGATCCGAGACATGTTCCACCTGCGGCTCTTGTTGAGACAGGACTCCGACGTGAGCTG 498
 DB 144 GluValArgAspLeuPheGlnIleMetLysLeuPheValAspThrAspAlaPheThrArgLeu 163
 QY 499 TTCGAGAGATTCTCCGGAGCGTG--CGCCGAGGAGGAGACCTGGAGCATTTGACG 555
 DB 164 SerArgArgValLeuArgAspIleSerGlnArgGlyArgAspLeuGlnIleLeuSer 183
 QY 556 CAGTACACACCTGTGTAAGCGCGCTTCGAGGAGTCTGCTGCCAAGAAAGTAT 615
 DB 184 GlnTyrIleThrPheValLysProAlaPheGlnGlnPheCysLeuProThrLysLysTyr 203
 QY 616 GCCGATGTGATCATCCAGAGAGTGCATATGTTGTCATCAACTGATCGGCGAG 675
 DB 204 AlaAspValIleIleProArgGlyAlaAspAsnLeuValAlaIleAsnLeuIleValGln 223
 QY 676 CACATCCAGACATTTCTGAATGTTGATCTGCAATGACCGACCGAGAGGTTCCAAATGG 735
 DB 224 HisIleGlnAspIleLeuAsnGly-----Gly 232
 QY 736 CGAGCTTAACCGACCTTTTCTGAGCGAGGAGCACCTCGGATGCTGACTTGAGC 795
 DB 233 ProSerLysArgGlnThr-----AsnGlyCysLeuAsnGlyTyrThrProSerArg 249
 QY 796 AAACGTCATTTGGAGTCCAG 831
 DB 250 LysArgGlnAlaSerGlnLysSerSerArgProHis 261
 RESULT 10
 ID AAB73494 standard; Protein; 261 AA.
 XX AAB73494;
 AC
 XX
 DT 31-JUL-2001 (first entry)
 XX
 DE Human transferase HTFS-1, SEQ ID NO:1.

XX	Human transferrase; HTRF; agonist; antagonist; cellular signalling;
KM	Proliferation; cell proliferative disorder; immune disorder;
KM	Atherosclerosis; hepatitis; psoriasis; cancer; tumour;
KM	Inflammation; AIDS; Addison's disease; allergy; ascuma; anaemia;
KM	Cirrhosis; Crohn's disease; atopic dermatitis; diabetes mellitus;
KM	Multiple sclerosis; rheumatoid arthritis; pancreatitis;
KM	Systemic lupus erythematosus; thrombocytopenia; ulcerative colitis;
KM	Hemodialysis; extracorporeal circulation; trauma; transgenic animal;
XX	gene therapy; drug screening.
OS	Homo sapiens.
XX	
XX	WO200132888-A2.
XX	
XX	10-MAY-2001.
XX	
XX	02-NOV-2000; 2000WO-US0485.
XX	
XX	04-NOV-1999; 99US-0163595.
XX	
XX	(INCY-) INCYTE GENOMICS INC.
XX	
PI	Tang YT, Yue H, Hillman JL, Lal P, Bandman O, Patterson C;
PI	Shih LL, Azimzai Y, Lu DM, Baughn MR;
XX	
XX	WPI; 2001-328796/34.
DR	N-PSDB; AAAH3801.
XX	
PT	Human transferrase polypeptides and polynucleotides useful for
PT	diagnosis, prevention and treatment of cell proliferative and immune
PT	system disorders and for identifying agonists and antagonists -
DS	Claim 1; Page 103-104; 157pp; English.

CC Sequence AAB73492-AAB73495 represent novel human transferase proteins
CC HRF5-1 to HRF5-42 and sequences AAH2801-AAH2842 represent cDNAs
CC encoding them. The proteins play important roles in the regulation of
CC cellular signalling and proliferation. The HRFs proteins are useful for
CC screening compounds for their effectiveness as agonists or antagonists of
CC transferase activity, or for compounds that specifically bind to an HRFs
CC protein or which modulates the activity of an HRFs protein.
CC Pharmaceutical compositions comprising an HRFs protein, HRFs
CC agonist or antagonist, or genetic construct encoding an HRFs
CC protein are useful for treating a disease or condition associated
CC with decreased or increased expression of functional HRFs. Disorders
CC which may be treated using such compositions include cell proliferative
CC disorders and immune disorders. For example, diseases which may be
CC treated include atherosclerosis, hepatitis, psoriasis, cancers (including
CC breast, bladder, bone marrow, brain and uterus cancer), inflammation,
CC AIDS, Addison's disease, allergies, asthma, anaemia, cirrhosis, Crohn's
CC disease, atopic dermatitis, diabetes mellitus, multiple sclerosis,
CC rheumatoid arthritis, pancreatitis, systemic lupus erythematosus,
CC thrombocytopenia, and ulcerative colitis. They may also be used to treat
CC complications of cancer, haemodialysis, extracorporeal circulation,
CC trauma and haematopoietic cancer, including lymphoma, leukaemia and
CC myeloma. Polynucleotides encoding HRFs proteins are useful for creating
CC transgenic animals to model human diseases, for diagnostic purposes and
CC to generate hybridisation probes useful in mapping the naturally
CC occurring genomic sequences. HRFs, and its catalytic or immunogenic
CC fragments are useful for screening libraries of compounds in a variety of
CC drug screening techniques. Antibodies which specifically bind HRFs may be
CC used for the diagnosis of disorders associated with the expression of
CC HRFs, or in assays to monitor patients being treated with HRFs or
CC agonists, antagonists or inhibitors of HRFs. The present sequence
CC represents an HRFs protein of the invention.

SQ Sequence 261 AA;

Alignment Scores:

Pred. No.: 1.66e-69

Score:

Percent Similarity: 79.418

Best Local Similarity: 69.49%

Query Match:	61.69%	Indels:	16
DB:	22	Gaps:	4

US-09-896-522-3 (1-834) X AAB73494 (1-261)

22 GACTGCGAGAGCCCCGCGCGGAGGCCGACCGTCCGCAC --- CAGCGCCCTTCTGATA 78

Db 4 AspSerGluInThrLeuGlnAsnHisGlnGlnProAsnGlyGlyGluProPheLeuIle 23

79 GGGTGAGCGCGGCACTGCCAGCGGAAGTCGACCGTGTGTGAGAAGATCATGGAGTTG 138

Db 24 GlyValSerGlyThrAlaSerGlyLysSerSerValCysAlaLysIleValGlnLeu 43

139 CTGGACAGACGAGGTGGAACAGCCGACGGAGGTGTATCCTGAGCCAGGACAGG 198

Db 44 LeuGlyGlnAsnGluValAspTyrArgGlnIleValGlnValIleLeuSerGlnAspSer 63

QY 199 TTCTACAAGTCTCTGACGGCAGAGCAGAAGCCCAAGGCTTGAAAGGACAGTACAATTT 258

Db 64 pHeTYrARgVAlLeuThrSerGIuGlnLysAlAlaLeuLysGIuGlnPheAsnPhe 83

259 GACCATCCAGATGCCCTTGATAATGATTGATGCACAGACTCTGAAGACATCGTGAG 318

Db 84 AspHisProAspAlaPheAspAsnGluLeuIleLeuLysThrLeuLysGluIleThrGlu 103

319 GGCAAAACGGTGGAGGTGCCGACCTATGATTTTGTGACACACTCAAGGTTACCAGAGACC 378

Db 104 GlyLysThrValGlnIleProValTyrAspPheValSerHisSerArgLysGluIuThr 123

379 ACGGTGCTACCCCTGCCGACGTGTTCTGTTGAGGGCATCTTGGTCTACAGCCAG 438

Db 124 ValThrValTyrProAlaAspValValLeuPheGluGlyIleLeuAlaPheTyrSerGln 143

439 GAGATCCGGACATGTTCCACCCTGGCCTCTTCGTGGACACCGACTCCGACGTCAGGCTG 498

Db 144 GluValArgAspLeuPheGlnMetLysLeuPheValAspThrArgLeu 163

499 TCTCGAAGAGTTCTCCGGGACGTG--CGCCGAGGGAGGACCTGGAGCAGATTCTGACG 555

Db 164 SerArgArgValLeuArgAspIleSerGluArgGlyArgAspLeuGlnIleLeuSer 183

556 CAGTACACCACCTTCGTGAAGCCGGCCTTCGAGGAGTCTGCCTGCCGACAAAGAAGTAT 611

Db 184 GlnTyrIleThrPheValLysProAlaPheGluGluPheCysLeuProThrLysLysTyr 203

616 GCCGATGTGATCATCCACGAGGAGTGCACAATATGTTGCCATCAACCTGATCGTCAG 675

Db 204 Alaaspvalllelleprowgilylaaspasnlleuvalalaileasnleuilevalgin 223

676 CACATCCAGGACATTCTGAATGGTGACATCTGCAATGGCACCAGGAGGTCCAATGGG 735

Db 224 HisIleGlnAspIleLeuAsnGly-----Gly 232

736 CGAGCTACAGCGACCTTTCTGAGCCAGGGGACCCTGGATGCTGACCTCTGGC 7955

Db 233 ProSerLysArgGlnThr-----AsnGlyCysLeuAsnGlyTyrThrProSerArg 2499

QY 796 AAACGGTCACATTTGGAGTCCAGCAGCAGACCCAC 831

Db 250 LysArgGlnAlaSerGluSerSerArgProHis 261

RESULT 11
ABD41303

ID ABP41393 standard; protein; 337 AA.

AC ABP41393;
XY

DT 22-AUG-2002 (first entry)

DE Human ovarian antigen HOELP29, SEQ ID NO:2525.

KW Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;
 KW ovarian cancer: breast cancer: tumour: reproductive system disorder:
 KW

KM infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
 KM PCOS; ovarian cyst; dysmenorrhea; endocrine disorder; infection;
 KM inflammatory condition; immune disorder; blood disorder; infection;
 KM cardiovascular disorder; respiratory disorder; neurological disorder;
 KM gastrointestinal disorder; urinary system disorder; drug screening;
 KM gene therapy; chromosome mapping; forensic analysis;
 KM antibody preparation; cytostatic; immunomodulatory; neuroprotective;
 KM antiinflammatory; gynaecological; reproductive; chromosome 1p32.
 XX
 OS Homo sapiens.
 PN WO20020677-A1.
 PD 03-JAN-2002.
 XX
 PF 07-JUN-2001; 2001WO-US18569.
 XX
 PR 07-JUN-2000; 2000US-209467P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Birse CE, Rosen CA;
 XX
 DR WPI; 2002-147878/19.
 XX
 DR N-PDB; ABQ54470.
 XX
 PT Isolated nucleic acid molecules encoding novel ovarian polypeptides,
 PT useful in the prevention, treatment and diagnosis of cancer (e.g.
 PT ovarian cancer), immune disorders, cardiovascular disorders and
 PT neurological diseases -
 XX
 PS Claim 11, SEQ ID NO 2525; 2922pp; English.
 XX
 CC The invention relates to 2175 novel human ovarian antigens (ABP41054-
 CC ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also
 CC encompasses polypeptides 90% identical and polynucleotides 95% identical
 CC to the sequences of the invention. The invention additionally relates to
 CC recombinant vectors and host cells comprising human ovarian antigen
 CC polynucleotides, antibodies against human ovarian antigens, and the use
 CC of ovarian antigen polynucleotides and polypeptides in diagnosing,
 CC treating, prognosing or preventing various ovary and/or breast-related
 CC disorders. Such conditions include ovarian cancer and breast cancer, and
 CC metastatic tumours of ovarian or breast origin, reproductive system
 CC disorders (e.g., infertility, disorders of pregnancy, anovulation,
 CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhea), endocrine
 CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic
 CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
 CC vaginitis), immune disorders (e.g., congenital and acquired
 CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
 CC blood-related disorders (e.g., anaemia), cardiovascular disorders,
 CC respiratory disorders, neurological disorders, gastrointestinal disorders
 CC and urinary system disorders. Ovarian antigen polypeptides and
 CC polynucleotides may also be used in screening for compounds which
 CC modulate ovarian antigen expression or activity. The polynucleotides may
 CC further be used for gene therapy, chromosome mapping, in the
 CC identification of individuals and in forensic analysis, and the
 CC polypeptides may be used as food additives or to prepare antibodies
 CC useful in disease diagnosis, drug targeting and phenotyping. The present
 CC sequence represents a human ovarian antigen of the invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WPIPO
 CC at ftp.wipo.int/pub/published_sequences.
 CC
 XX
 SQ Sequence 337 AA;

Alignment Scores:
 Pred. No.: 7.62e-67 Length: 337
 Score: 902.50 Matches: 192
 Percent Similarity: 73.65% Conservative: 26
 Best Local Similarity: 64.86% Mismatches: 44
 Query Match: 59.61% Indels: 34
 DB: 23 Gaps: 6

US-09-896-522-3 (1-834) x ABP41393 (1-337)
 QY 3 GGCTTGGGGGGAGGGA-----AGACTGCGAGAGCCCGCGGAGGCGCA--- 50
 DB 57 GYAlaGlyGlyGlnArgAlaSerValArgThrGlySerGlyArgArgGlyGlyAlaAsn 76
 QY 51 -----CCGTCGCGACCA-----GGCGCC 68
 DB 77 HisGlyArgGlyGlnArgAlaSerProAlaGluProProAlaAlaGlnArgArgAla 96
 QY 69 CTTCTGATAGGGGTGAGCGCGGCACTGCGAGCGGGAATCGACCTGTGTGAGAAGAT 128
 DB 97 LeuProTyrArgArgHis-GlyGlyThrAlaSerGlyLysSerSerValCysAlaLys 116
 QY 129 CATGAGTTGCTGGGACAGACAGAGGTGGAACAGCGGACCGGAAGTGTCTCTGAG 188
 DB 116 eValGlnLeuLeuGlyGlnAsnGlnValAspTyrArgGlnLysGlnValAlaLeu 136
 QY 189 CCAGACAGGTTCTCAAGATCCTGACGCGAGAGGAGGAGCCCTTGAAAGACA 248
 DB 136 rGlnAspSerPheTyrArgValLeuThrSerGlnLysAlaValAlaLeuLys**GI 156
 QY 249 GTACAAATTTGACCATCCAGATGCCCTTGATTAATGATTTGATGACAGACTGGAAGA 308
 DB 156 nPheAsnPheAspHisProAspAlaPheAspAsnGlnLeuLeuLeuThrLeuLysGI 176
 QY 309 CATCGTGGAGGGGCAAAACGCTGAGAGTCCGACCTATGATTTTGTACACACTCAAGCTT 368
 DB 176 uIleThrGlnGlnLysThrValGlnLeuProValTyrAspPheValSerHisSerArg 196
 QY 369 ACCAGACGACGCGGTCTCAACCTCGGAGCGGAGTCTGTGAGGCACTGTGAGTT 428
 DB 196 sGlnGlnThrValThrValTyrProAlaAspValValLeuPheGlnGlyLeuAla 216
 QY 429 CTACAGCAGAGATCCGAGGACATGTTCCACCTGCGCCTTTCGTGACACCGACTCGA 488
 DB 216 eTyrSerGlnGlnValArgAspLeuPheGlnMetLysLeuPheValAspThrAspAla 236
 QY 489 CGTCAGGCTGTCTCGAAGAGTTCTCCGGACGTG---CGCCGAGGAGGAGCCTGGAGCA 545
 DB 236 pThrArgLeuSerArgArgValLeuArgAspLieserGlnArgGlyArgAspLeuGln 256
 QY 546 GATTGAGCAGATACACACCTGCGAAGCGGCTTCGAGGAGTTCGCTCGCGAC 605
 DB 256 nIleLeuSerGlnTyrIleThrPheValLysProAlaPheGlnGlnPheCysLeuPro 276
 QY 606 AAGAAGTATGCGATGATGATCCACAGAGAGTGAACAATATGTTGCCATCAACT 665
 DB 276 rLysLysTyrAlaAspValIleIleProArgGlyAlaAspAsnLeuValAlaIleAsn 296
 QY 666 GATCGTCAGACATCCAGACCATTTGATGTTGACATCTGCAATGGCACCGAGAGG 725
 DB 296 uIleValGlnHisIleGlnAspLleLeuAsnGly----- 307
 QY 726 GTCCATATGGGCGAGTACAGGACCTTTCTGAGCCAGGAGACCACTGGAGTGT 785
 DB 308 -----GlyProSerLysArgGlnThr-----AsnGlyCysLeuAsnGlyTyr 322
 QY 786 GACCTCTGCAACAGGTCAATTGAGATCGACAGACAGACCCAC 831
 DB 322 rProSerArgLysArgGlnAlaSerGlnSerSerArgProHis 337
 RESULT 12
 AAB56582
 ID AAB56582 standard; Protein; 337 AA.
 XX
 AC AAB56582;
 XX
 DT 13-MAR-2001 (first entry)
 XX Human prostate cancer antigen protein sequence SEQ ID NO:1160.
 DE Human; prostate cancer; prostate cancer antigen; detection; diagnosis;
 XX

[illegible]

QY	169	CGAGGACAGGCTCTCAAGAGTCCCTGACGGCAGACGAGGACCGCAAGCCCTTGAAGAACA	248
Db	136	rglnaspserPheItyrArgValIleuthSerIuglnuSAlaLysAlaLeuLys**G	156
QY	249	GTACAAATTTTGCACATCCAGATGCGCTTGATATGATTTGATGACACAGACTTGAAGAA	308
Db	156	nPhenalnPhenaphsPronaspAlaPheaspnngnu***IleuLysThrIleuLysgl	176
QY	309	CATCGTGAAGGGCAAAACGGTGAAGGTGCGACCTATGATTTTGTGACACTCAAGTT	368
Db	176	uIleThrIugIugIyLsthrValGlnIleProValTyrAspPheValSerHisSerArgLy	196
QY	359	ACCAAGACACACAGGTGCTTACCCCTGGCGGACGTGGTCTGTTGAGGGATCTGTGTT	428
Db	196	sgIuIuIthrValIthrValTyrProAlaaspValIleuPheIugIuYIleuAlaPh	216
QY	429	CTACAGCCAGAGATCCGGGACATGTTCCACCTGGCGCCTCTTGAGAACCGACTCCGA	488
Db	216	eIyrsTerIugIuValAlaArgaspLeuPheIuIleuTylsLeuPheValAspThrAspAlaAs	236
QY	489	CGTCAGGCTGTCTCGAAGAGTTCCTCCGGGACGTG--CGCCGAGGGAGGACCTGAGCA	545
Db	236	pThrArgLeuSerAArgValIleuArgaspIleSerGluArgGlyArgAspIeuIugI	256
QY	546	GATTCTGACGACGATACACCACTTGTGTAAGCCGGCCTTCCAGAGACTTCTGCTGCCGAC	605
Db	256	nIleLeuSerGlnTyrIleThrPheValIlyspProAlaPheIugIuPheCysIeuProTh	276
QY	606	AAAGAAGATGCGCGATGTGATCATCTCCACGAGAGGTGACAAATATAGTGGCCATCAACT	665
Db	276	flYalysItyrAlaaspValIleIleProArgGlyAlaaspAluIleValAlaIleasne	296
QY	666	GATCGTCGACGACATCCAGAGACATTTCTGAATGGTGAACATCTCAATGGCACCGAGAGG	725
Db	296	uIleValGlnHisIleGlnAspIleLeuasnIy-----	307
QY	726	GTCCCAATGGCGGAGACTACAGCGGACCTTTTCTAGACGGGAGCACACCTGGGATGCT	785
Db	308	-----GlyProSerIyAsrGlnThr-----AsnGlyCylLeuasnIyTyrtH	322
QY	786	GACCTCTGGCAACGGTCACATTTGAGTGCAGACGAGACCCGAC	831
Db	322	rProSerArgIyAsrGlnAlaSerGlnSerSerSerThrArgProHis	337
RESULT 13			
AAM40480			
ID	AAM40480 standard; Protein; 335 AA.		
XX			
AC	AAM40480;		
XX			
DT	22-OCT-2001 (first entry)		
XX			
DE	Human polypeptide SEQ ID NO 5411.		
XX			
KM	Human; noctropic; immunosuppressant; cytostatic; gene therapy; cancer;		
KM	peripheral nervous system; neuropathy; central nervous system; CNS;		
KM	Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;		
KM	amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;		
KM	chemokinetic; thrombolytic; drug screening; arthritis; inflammation;		
KM	leukemia.		
XX			
OS	Homo sapiens.		
XX			
PN	WO200153312-A1.		
XX			
PD	26-JUL-2001.		
XX			
PF	26-DEC-2000; 2000WO-US34263.		
XX			
PR	21-JAN-2000; 2000US-0486725.		
PR	25-APR-2000; 2000US-0552317.		
PR	09-JUL-2000; 2000US-0598042.		
PR	19-JUL-2000; 2000US-0620312.		

PR	03- AUG-2000; 2000US-0653450.
PR	14-SEP-2000; 2000US-0662191.
PR	19-OCT-2000; 2000US-0693036.
PR	29-NOV-2000; 2000US-0727344.
XX	
PA	(HXSE-) HXSEQ INC.
XX	
PI	Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI	Wang J, Wang Z, Weinman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI	Zhao QA, Zhou P, Goodrich R, Drmanac RJ;
XX	
DR	WPI; 2001-442253/47.
DR	N-PSDB; AAI59636.
XX	
PT	Novel nucleic acids and polypeptides, useful for treating disorders
PT	such as central nervous system injuries -
XX	
PS	Example 2; SEQ ID NO 5411; 10078bp; English.
XX	
CC	The invention relates to human nucleic acids (AA157798-AA161369) and
CC	the encoded polypeptides (AA138642-AA442213) with neurotrophic,
CC	immunosuppressant and cytosaratic activity. The polynucleotides are useful
CC	in gene therapy. A composition containing a polypeptide or polynucleotide
CC	of the invention may be used to treat diseases of the peripheral nervous
CC	system, such as peripheral nervous injuries, peripheral neuropathy and
CC	Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC	lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC	utilization of the activities such as: Immune system suppression,
CC	Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC	and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC	assays for receptor activity, arthritis and inflammation, leukaemias and
CC	C.N.S disorders.
CC	Note: The sequence data for this patent did not form part of the printed
CC	specification.
XX	
SO	Sequence 335 AA;
XX	
Alignment Scores:	
Pred. No.:	1,02e-65 Length: 335
Score:	889.00 Matches: 190
Percent Similarity:	72.488 Conservative: 26
Best Local Similarity:	63.768 Mismatches: 46
Query Match:	58.724 Indels: 36
DB:	22 Gaps: 7
US-09-896-522-3 (1-834) x AAM40480 (1-335)	
OY	4 GCTTCGCGCGGAGCGCA-----21
DB	52 AAGIYAAAGIAGIYGIUnrGlaIaserValArgThrGIYserGIYArGArGIYArg 71
OY	22 -----GACTGCGAGAGCCCCGCGCGGAGCCGACCGTCCGAC---CAGCG 66
DB	72 ThrMetAlaGlyAAspSerGIUnrIleuGlnAsnHisGlnInProAsnGIYGIYGIUn 91
OY	67 CCCTTCCTGATYGGGGGTGAGCGCGCGGACCTGCGAGCGGGAAGTCAACCGGTGGAAG 126
DB	92 ProPheIleuIleGIYAlaSerGIYGIYThrAlaSerGIYIYSerSerValCYsAlaIYs 111
OY	127 ATCATGAGACTTCGTGGAGCAGAAcGAGGTGGAACAGCGCGGAGGAAAGGTGATCTCTG 186
DB	112 ILeValGlnIleuLeuGIYGIUnenGIYValaAspTYArGIUnIYsGIUnValIleLeu 131
OY	187 AGCCGAGAGAGGTTCTTACAAGGTCTCTGACGGCAGAGCAGAAAGGCCAAAGGCTTGAAAGA 246
DB	132 SerGIAspSerPheTYArGIYValIleuTrISerGIUnIYsAlaIYsAlaLeuYsGIY 151
OY	247 CAGTACAAATTTTGACCATCGAGATGCTTGTGATTAATGATTGATGACAGAGCTTGAAG 306
DB	152 GlnPheAsnPheAspHisProAspAlaIaPheAspAsnGIUnIleuIYsTrIleuYs 171
OY	307 AACATCGTGAAGGCAAAACGCTGAGAGTGCAGCCTATGATTTTGTGACACCTCAAG 366

Db	172	GIuIleThrGIuGIuLyLyThrValGlnIleProValTyrAspPheValSerHisSerArg	191
Qy	367	TTACCAAGAACCCAGCGTGTCTACCCCTCGGACGTCGTCTTTGAGGGCATCTTGGTG	426
Db	192	LySGluGIuIuThrValThrValTyrProAlaAspValValIleuPheGluIleuAla	211
Qy	427	TTCTTACAGCCAGGAG--ATCCGGGACATGTTCCACCTCGCGCTCTTCTGTGACACCCGAC	483
Db	212	PheTyrSerGIuGIuArgIleArgAspIleuPheGlnMetLysIleuPheValAspThrAsp	231
Qy	484	TCCGACGTCAGGCTGTCTCGAAGATTCTCCGGACGTC--CGCCAGAGGAGGACCTG	540
Db	232	AlaAspThrArgIleuSerArgIleuValIleuLysAspIleSerGIuArgIleLysAspIleu	251
Qy	541	GAGCAGATTCTGACGCGACGATACAC--ACCTTCGTAAACCGGCTTCGAGGATTTCTGC	597
Db	252	GIuGIuIleIleuSerSerSerThrIleuArgPheValLysProAlaPheGIuIuPheCys	271
Qy	598	CTGCCGACAAAGAGATGATCCGATGTGATCATCCACGAGAGTGACAAATATGTTGCC	657
Db	272	LeuProProlLysTyrAlaAspValIleIleProArgGIuAlaAspAsnArgValPro	291
Qy	658	ATCAACCTCGATCGGCAGACATCCAGACATTCGATGTCGATGTCGACATCTGCAATGGCAC	717
Db	292	IleAsnIleuIleValGlnHisIleGlnAspIleIleuAsnGly-----	305
Qy	718	CGAGGAGGTCCAATGAGCGGACGTACAAAGCGGACCTTTCTGAGCCAGGAGCACCTT	777
Db	306	-----GlyProSerAsnArgIuIuIu-----AsnGlyCysIleuAsn	317
Qy	778	GCGATGCTGACCTCTGGCAAAACGTCACATTGAGTCCAGCAGACGACCCAC	831
Db	318	GlyTyrThrProSerArgLysArgIuAlaSerGIuSerSerArgProHis	335

RESULT 14

ABB62307

ID ABB62307 standard; Protein; 260 AA.

AC ABB62307;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 13713.

XX Drosophila: developmental biology; cell signalling; insecticide;

KW pharmaceutical.

OS Drosophila melanogaster.

OS WO200171042-A2.

PN 27-SEP-2001.

PD 27-SEP-2001.

PF 23-MAR-2001; 2001WO-US09231.

XX 23-MAR-2000; 2000US-191637P.

PR 11-JUL-2000; 2000US-0614150.

XX (PEKE) PE CORP NY.

PA Venter JC, Adams M, Li PWD, Myers EW;

PI WPI; 2001-656860/75.

DR N-P8DB; ABL06410.

XX New isolated nucleic acid detection reagent for detecting 1000 or more

PT genes from Drosophila and for elucidating cell signalling and cell-cell

PT interactions -

XX

PS Disclosure; SEQ ID NO 13713; 21pp + Sequence listing; English.

XX

CC The invention relates to an isolated nucleic acid detection reagent

capable of detecting 1000 or more genes from *Drosophila*. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABU16176-ABU30511), expressed DNA sequences (ABU57737-ABU72072).

The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

Sequence 260 AA:

Alignment Scores:

Pred. No.:	3,73e-53	Length:	260
Score:	738.00	Matches:	143
Percent Similarity:	82.94%	Conservative:	32
Best Local Similarity:	67.77%	Mismatches:	34
Query Match:	48.75%	Indels:	2
DB:	22	Gaps:	2

US-09-896-522-3 (1-834) x ABG27216 (1-260)

QY 67 CCTTCTGATAGGAGGTGAGCGCGGCACTGCCAGCGGAAGTCCAGCCGTGTGAGAAG 126
 |||||
 Db 27 ProPheLeuIleGlyValAlaGlyGlyThraIaSerGlySerThrValCysLeuLys 46
 |||||
 QY 127 ATCATGAGTTCGTCGGGACAGACAGAGTGAACAGCGCGGCGGAAGTGTGATCCTG 186
 |||||
 Db 47 IleMetGluIleuGlyGlnAlaGluMetAspHisThrGlnArgIleValSerIle 66
 |||||
 QY 187 AGCCGAGCAGAGTTCACAGAGTCTGACGACAGACAGACAGACAGCGCTTGAAGA 246
 |||||
 Db 67 SerGlnAspSerPheIleYrArgGluIleuThrProAlaGluLysAlaIleGlnLysGly 86
 |||||
 QY 247 CAGTACAAATTTGACATCCAGATGCTTGAATATGATTTGATGACAGAGCTTGAAG 306
 |||||
 Db 87 LeuPheAsnPheAspHisProAspAlaPheAsnGluGluMetCysSerThrLeuGln 106
 |||||
 QY 307 AACCTCGAGGCGCAAAAGCGTGGAGGTCCGACCTATGATTTTGTGACACTCA--- 363
 |||||
 Db 107 AsnIleuLysGlyHisLysValGluIleProSerIleYrAspIleYrAsnSerLeu 126
 |||||
 QY 364 AGGTACAGAGACACAGCGTGTACCTCGGACAGTGTTCACCTGCGCTTGTGACACCGAC 423
 |||||
 Db 127 AspPheGluAsnValIleuValIleIleYrProAlaAspValValIleuPheGluGlyIleLeu 146
 |||||
 QY 424 GTGTTCTACAGCCAGAGATCCGGAACATGTTCCACCTGCGCTTGTGACACCGAC 483
 |||||
 Db 147 ValPheIleYrPheProLysIleArgGluIleuPheHisMetLysLeuPheValAspThrAsp 166
 |||||
 QY 484 TCCGACGTCGAGCGTCTCGAAGAGTCTCCGAGACGCG---CGCCGAGGAGGAGACTG 540
 |||||
 Db 167 SerAspThrArgIleuAlaArgValProArgAspIleAsnGluArgGlyArgAspLeu 186
 |||||
 QY 541 GAGCAGATTCTGACGACGATACACACCTTCGTGAAGCGGCTTCGAGAGTTGTGCTG 600
 |||||
 Db 187 AspIleValIleuThrGlnIleYrMetThrPheValLysProAlaPheGluGluPheCysSer 206
 |||||
 QY 601 CCGGCAAGAAGTATGCCGATGTATCATCCACGAGAGTGAACAATATGTTGCCATC 660
 |||||
 Db 207 ProThrLysLysPheAlaAspValIleIleProArgGlyAlaAspAsnThrValAlaIle 226
 |||||
 QY 661 AACCTGATCGTCGACGATCCAGACACTTGTG 693
 |||||
 Db 227 AspLeuIleValHisHisIleGlyGluIleLeu 237
 |||||

18-FEB-2002 (first entry)

DE Novel human diagnostic protein #27207.

KW Human, chromosome mapping; gene mapping; gene therapy; forensic; food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

PN WO200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US08631.

PR 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

PA (HYSE-) HYSEQ INC.

PI Dermanac RT, Liu C, Tang YT.

DR WPI, 2001-639362/73.

DR N-PSDB; AAS91403.

XX New isolated polynucleotide and encoded polypeptides, useful in PT diagnosis, forensic, gene mapping, identification of mutations PT responsible for genetic disorders or other traits and to assess PT biodiversity.

XX Claim 20; SEQ ID No 57575; 103bp; English.

XX The invention relates to isolated polynucleotide (I) and CC polypeptide (II) sequences. (II) is useful as hybridisation probe, CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome CC and gene mapping, and in recombinant production of (II). The CC polynucleotides are also used in diagnostics as expressed sequence tags CC for identifying expressed genes. (I) is useful in gene therapy techniques CC to restore normal activity of (II) or to treat disease states involving CC (II). (II) is useful for generating antibodies against it, detecting or CC quantitating a polypeptide in tissue, as molecular weight markers and as CC a food supplement. (II) and its binding partners are useful in medical CC imaging of sites expressing (II). (I) and (II) are useful for treating CC disorders involving aberrant protein expression or biological activity. CC The polypeptide and polynucleotide sequences have applications in CC diagnostics, forensics, gene mapping, identification of mutations CC responsible for genetic disorders or other traits to assess biodiversity CC and to produce other types of data and products dependent on DNA and CC amino acid sequences. ABG0010-ABG30377 represent novel human CC diagnostic amino acid sequences of the invention. CC Note: The sequence data for this patent did not appear in the printed CC specification, but was obtained in electronic format directly from WIPO CC at ftp.wipo.int/pub/published_pct_sequences.

XX

Sequence 326 AA:

Alignment Scores:

Pred. No.:	1.08e-40	Length:	326
Score:	589.00	Matches:	132
Percent Similarity:	63.64%	Conservative:	29
Best Local Similarity:	52.17%	Mismatches:	52
Query Match:	38.90%	Indels:	40
DB:	22	Gaps:	7

US-09-896-522-3 (1-834) x ABG27216 (1-326)

QY 4 GCTTCGCGGAGGCGAGAGCTGCGAGACCCCGCGCG---GAGCGCAGCGTCCGCAC 60
 |||||
 Db 67 AlaaGlnAlaAlaGlyGlyGlyAlaGluProTrpProGlyThrIaSerThrProCys 86
 |||||
 QY 61 CAG-----CGGCCCTTCTGATAGGAGGTGACGCGCGCACT 96
 |||||
 Db 87 ArgThrThrSerSerProThrAlaAlaSerThrSerLeuIleGlyValThrTrpGlyThr 106
 |||||

```
Qy 97 GCCAGCGGGAAGTGCACCGTGTGTGAGAGATCATGGAGTTGCTGGGACAGAAAGAGGTG 156
Db 107 AlaSerGlyLysSerSerValCysAlaLysIleValGlnLeuLeuGlyGlnAsnGlnVal 126
Qy 157 GAAACAGCGGACGCGAGAGTGTGCATCTGTGACCGAGACAGATTCTACAAGTCTTGACG 216
Db 127 AspTyrArgGlnLysGlnValValIleLeuSerGlnAspSerPheTyrArgValLeuThr 146
Qy 217 GCAGAGCAGAAAGCGCAAGGCTTGAANGACAGTACAATTGACCATTCAGATGCTTT 276
Db 147 SerGlnGlnLysAlaLysAlaLeuLysGlyGlnPheAsnPheAspHisProAspAlaPhe 166
Qy 277 GATPATGATTGTTGACACAGACTCTGAAGACATCGTGAGGCGCAAAACGTGAGAGTG 336
Db 167 AspAsnGlnLeuIleLeuLysThrLeuLysGlnIleThrGlnGlyLysThrValGlnIle 186
Qy 337 CCGACTATGATTTTGTGACACACTCAAGTTACAGAGACACACGCGTGTCTACCTGCG 396
Db 187 ProValTyrAspPheValSerHisSerArgLysGlnIleThrValThrValTyrProAla 206
Qy 397 GACGTGTTCTGTTTGAAGGCATCTTGTGTTCTACAGCCAGAGATCCGGACATGTT 456
Db 207 AspValValLeuPheGlnGlyIleLeuAlaPheTyrSerGlnGlnValArgAspLeuPhe 226
Qy 457 CACCTGGCCTCTTGTGACACCGACTCCGAGCTCAGGCTGTCTGAGAGA---GTTCTC 513
Db 227 GlnMetLysLeuPheValAspThrAspAlaAspThrArgLeuSerArgArgGlyIleMet 246
Qy 514 CCGGACGTGCGCCGAGGAGGACCTTGAGCAGATT----- 549
Db 247 AsnLeuLeuLeuHisProArgGlyLeuArgAlaIleThrIleAlaValPheGlyLysGln 266
Qy 550 -----CTGACGCAGTACACCACTTCGTGAAGCCGCC----- 582
Db 267 AsnThrTyrIleArgLeuGlnProPheArgIleAsnValProProThrIleThrLysHis 286
Qy 583 TTCGAGAGATTCT---TGCTGCGCG----- 603
Db 287 IleGlnIleLeuGlnCysAlaProValValThrGlnLeuSerArgLysGlnIleThrVal 306
Qy 604 ACAAGAAGTATGCCGATGTGATCATCCACGAGAGAGTG 642
Db 307 ThrValTyrProAlaAspValValLeuPheGlnGlyIle 319
```

Search completed: November 25, 2003, 07:30:45
Job time : 44.5818 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2003, 07:58:14 ; Search time 34 Seconds
(without alignments)
2102.370 Million cell updates/sec

Title: US-09-896-522-2

Sequence: 1 MASAGEDCSPAPADRP.....HPGMLTSGKSHLESSSRPH 277

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

.SPTREMBL_23:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriophage:*
17: sp_archaea:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	932	64.3	261	11	Q8C476
2	873	60.2	201	4	Q96B00
3	733	59.6	260	5	Q8MRJ1
4	570.5	39.3	486	10	Q9FKS0
5	568	39.2	465	10	Q9LTV6
6	563	38.8	483	10	Q9LTK4
7	559	38.6	515	5	Q19583
8	559	38.6	555	5	Q9U317
9	544	37.5	466	10	Q8VYB2
10	542	37.4	614	5	Q9V610
11	540.5	37.3	626	5	Q8MOK4
12	529	36.5	419	4	Q8N524
13	507.5	35.0	469	10	Q8LD95
14	506	34.9	542	10	Q9L632
15	459	31.7	209	16	Q9LFF2
16					Q8E0A8

17	459	31.7	454	3	Q74427	074427 schizosacch
18	458	31.6	209	16	Q8E5Y5	Q8E5Y5 streptococc
19	451	31.1	207	16	Q8CSB2	Q8CSB2 streptococc
20	441	30.4	209	16	Q8DPT1	Q8DPT1 streptococc
21	439	30.3	419	5	Q9BMX4	Q9BMX4 cryptospori
22	404	27.9	210	16	Q8BPT5	Q8BPT5 oceanobacti
23	403	27.8	213	16	Q8BPF4	Q8BPF4 vibrio vuln
24	390.5	26.9	111	4	Q9BU42	Q9BU42 homo sapien
25	380.5	26.2	111	4	Q92528	Q92528 homo sapien
26	369	25.4	105	11	Q9QY68	Q9QY68 rattus norv
27	349	24.1	212	16	Q8EDX4	Q8EDX4 shewanella
28	233	16.1	674	10	Q9C664	Q9C664 arabidopsis
29	212	14.6	643	10	Q9C9B9	Q9C9B9 arabidopsis
30	210.5	14.5	334	16	Q8YPR9	Q8YPR9 anabaena sp
31	210.5	14.5	403	10	Q8GRU9	Q8GRU9 oryza sativ
32	207	14.3	646	10	Q8S3R9	Q8S3R9 oryza sativ
33	206.5	14.2	396	10	Q9S033	Q9S033 odontella s
34	200	13.8	334	16	Q8L0S5	Q8L0S5 synecchococ
35	200	13.8	334	16	Q8DHN2	Q8DHN2 synecchococ
36	199.5	13.8	352	10	P93681	P93681 pium sativ
37	197	13.6	367	3	Q94642	Q94642 schizosacch
38	196.5	13.6	448	10	Q8GUE1	Q8GUE1 galidieria s
39	195.5	13.5	333	2	Q9L8V7	Q9L8V7 synecchococ
40	191	13.2	313	16	Q8YUJ6	Q8YUJ6 anabaena sp
41	187.5	12.9	271	10	Q9SPH7	Q9SPH7 beta vulgar
42	178.5	12.3	232	3	Q12084	Q12084 saccharomyc
43	172	11.9	555	16	Q9WZL0	Q9WZL0 thermocoga
44	169.5	11.7	405	10	Q9ATC3	Q9ATC3 vaucheria l
45	156.5	10.8	307	16	Q8DD30	Q8DD30 vibrio vuln

ALIGNMENTS

RESULT 1
ID Q8C476 PRELIMINARY; PRT; 261 AA.

AC Q8C476;
DT 01-MAR-2003 (T-EMBLrel. 23, Created)
DT 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)
DE Uridine-cytidine kinase 2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RX MEDLINE=22354683; PubMed=12466651;
RA The FANTOM Consortium
RA The RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573 (2002).
DR EMBL; AK082837; BAC38646.1; -
SQ SEQUENCE 261 AA; 29438 MW; 780AA3C298A8153 CRC64;

Query Match 64.3%; Score 932; DB 11; Length 261;
Best Local Similarity 68.8%; Pred. No. 1.4e-75;
Matches 187; Conservative 28; Mismatches 41; Indels 16; Gaps 4;
QY 8 DCSPAPADRP...QRPFLIGVSGTSGKSTVCEKIMELQNEVEORQKRVIIISQR 66
DB 4 DSEQTLQNHQDPNGEPFLIGVSGTSGKSTVCAKIQLQNEVDYHQKQVILISQDS 63
QY 67 FKVVLTAQKAKALNGOVNFDPDAFNDIMHRTKNIYEGKTVETPVDFTHSRPT 126
DB 64 FKVLTSEQAKALNGOVNFDPDAFNDIMHRTKNIYEGKTVETPVDFTHSRPT 123
QY 127 TVVYPADVVLFEGLIVFYSQEIIRDMFHLRLFVDTSDVRLSRVRVADV...RRGRDLFQILT 185
DB 124 VVIYPADVVLFEGLIVFYSQEIIRDMFHLRLFVDTSDVRLSRVRVADV...RRGRDLFQILT 183

QY 186 QYTFVYKPAFEFCLPTKKYADVILPRGVDMVAIINLVQHIDILNGDICKHRRGSSNG 245
 DB 184 QYIFVVPAPAFEFCLPTKKYADVILPRGADVLVAIINLVQHIDILNGDICKHRRGSSNG 232
 QY 246 RSYKRTFSEPDHGMILTGKRSKLSSESRPH 277
 DB 233 LSKRQT---NGYFNGYPSRRQASSESSSRPH 261

RESULT 2

Q96BJ0 PRELIMINARY; PRT; 201 AA.
 AC Q96BJ0: 01-DEC-2001 (Tremblrel. 19, Created)
 DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
 DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
 DE Similar to uridine-cytidine kinase 1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Colon;
 RA Strauberg R.;
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC015547; AAH1547.1; -.
 DR InterPro; IPR006083; PRK URK.
 DR InterPro; IPR000764; Uridine_kin.
 DR Pfam; PF00485; PRK; 1.
 DR PRINTS; PR00986; URIDINKINASE.
 KM Kinase.
 SQ SEQUENCE 201 AA; 22790 MW; 0B5F2F0FF7B5363 CRC64;

Query Match 60.2%; Score 873; DB 4; Length 201;
 Best Local Similarity 100.0%; Pred. No. 2e-70;
 Matches 169; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASAGGEDCSAPADRPHPORPPLIGVSGGTASGKSTVCEKIMELLGQNEVEQORRV 60
 DB 1 MASAGGEDCSAPADRPHPORPPLIGVSGGTASGKSTVCEKIMELLGQNEVEQORRV 60
 QY 61 ILSODRFKVLTAOKAKALKGQVNDHPDAPDNDLMHRTLKNVGEKTVVPTDYVTH 120
 DB 61 ILSODRFKVLTAOKAKALKGQVNDHPDAPDNDLMHRTLKNVGEKTVVPTDYVTH 120
 QY 121 SRLPETHVVPADVLPFEGILVFYSOEIRDMFHLRLFVDTDSVRLSR 169
 DB 121 SRLPETHVVPADVLPFEGILVFYSOEIRDMFHLRLFVDTDSVRLSR 169

RESULT 3

Q8MRJ1 PRELIMINARY; PRT; 260 AA.
 AC Q8MRJ1: 01-OCT-2002 (Tremblrel. 22, Created)
 DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)
 DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
 DE LDI3909P.
 GN CG6364.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Archipoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkley;
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Dresnak D., Farfan D., Fiske E.,
 RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
 RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,

RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.B., Rubin G.M.,
 RA Ceiniker S.; (JUN-2002) to the EMBL/GenBank/DBJ databases.
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY119583; AAM50237.1; -.
 DR FlyBase; FBgn0039179; CG6364.
 DR InterPro; IPR006083; PRK URK.
 DR InterPro; IPR000764; Uridine_kin.
 DR Pfam; PF00485; PRK; 1.
 DR PRINTS; PR00986; URIDINKINASE.
 SQ SEQUENCE 260 AA; 29223 MW; 03CAA6DCA04A1E5 CRC64;

Query Match 50.6%; Score 733; DB 5; Length 260;
 Best Local Similarity 67.3%; Pred. No. 1.e-57;
 Matches 142; Conservative 32; Mismatches 35; Indels 2; Gaps 2;

QY 23 PFLIGVSGGTASGKSTVCEKIMELLGQNEVEQORRVYILSDRFKVLTAOKAKALG 82
 DB 27 PFLIGVAGGTASGKSTVCKKIMEQLQAEWDHTRQVSVISQDSFYRELTPEAKAKAKG 86
 QY 83 QVNFDPDAPDNDLMHRTLKNVGEKTVVPTDYVTHS-RLEPETHVVPADVLPFEGIL 141
 DB 87 LFNPDHPDAPNEELMTSLQNLKSHKVELPSYDRTNSLDPEENVLYIPADVLPFEGIL 146
 QY 142 VFYSOEIRDMFHLRLFVDTDSVRLSRVLDVY-RRGRDLEQILTYTFVYKPAFEFCL 200
 DB 147 VFYFPRKIRELFHMKLTFVDTDPDTRLARVPDINERGRDLDAVLTYGMYFVKPAFEFCS 206
 QY 201 PTKYADVILPRGVDMVAIINLVQHIDIL 231
 DB 207 PTKKPADVILPRGADVLVAIINLVQHIGETL 237

RESULT 4

Q9FKS0 PRELIMINARY; PRT; 486 AA.
 AC Q9FKS0: 01-MAR-2001 (Tremblrel. 16, Created)
 DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
 DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
 DE Uridine kinase-like protein.
 GN AT5G40870.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Colombia;
 RX MEDLINE=98344145; PubMed=9679202;
 RA Kaneko T., Kotani H., Nakamura Y., Sato S., Asamizu E., Miyajima N.,
 RA Tabata S.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. V. Sequence
 RT features of the regions of 1,381,565 bp covered by twenty one
 RT physically assigned pl and TAC clones.";
 RL DNA Res. 5:131-145(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Yamada K., Chan M.M., Chang C.H., Dale J.M., Hsuan V.W., Lee J.M.,
 RA Onodera C.S., Quach H.L., Tang C.C., Toriumi M., Wong C., Wu H.C.,
 RA Yu G., Yuan S., Chen H., Cheuk R., Jones T., Kim C.J., Nguyen M.,
 RA Palm C.J., Shin P., Southwick A., Tripp M.G., Wu T., Davis R.W.,
 RA Ecker J.R., Theologis A.;
 RT "Arabidopsis Open Reading Frame (ORF) Clones.";
 RT Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
 RL EMBL; AB011477; BAB11349.1; -.
 DR EMBL; BT002336; AAN86169.1; -.
 DR HSSP; Q26998; 1BD3.
 DR InterPro; IPR006083; PRK URK.
 DR InterPro; IPR000836; PRTtransferase.
 DR InterPro; IPR000764; Uridine_kin.
 DR Pfam; PF00156; Pribosyltran; 1.
 DR Pfam; PF00485; PRK; 1.

Db 229 DFLVPSKKYADVIFPRGGDNHVAVDLIVQHTLGGHDLCKY---PNVFIETTTQIR 285
 QY 256 GDH 258
 Db 286 GMD 288

RESULT 7

Q19583 PRELIMINARY; PRT; 515 AA.
 AC Q19583;
 DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
 DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)
 DE F19B6.1a protein.
 GN F19B6.1 OR F19B6.1A.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 CX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Thomas K.;
 RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99069613; PubMed=9851916;
 RA none;
 RT "Genome sequence of the nematode C. elegans: A platform for
 RT investigating biology";
 RL Science 282:2012-2018(1998).
 DR EMBL; Z69635; CA93459.1; -.
 DR HSSP; Q26998; 1BD3.
 DR WormPep; F19B6.1a; CE05666.
 DR InterPro; IPR006082; PRK.
 DR InterPro; IPR000764; Uridine_kin.
 DR Pfam; PF00485; PRK.1.
 DR PRINTS; PR00478; PHRIBLKINASE.
 DR PRINTS; PR00988; URIDINKINASE.
 DR TIGRFAMs; TIGR00235; udk; 1.
 SQ SEQUENCE 515 AA; 58452 MW; 9584D947A2D3B5C5 CRC64;

Query Match 38.6%; Score 559; DB 5; Length 515;
 Best Local Similarity 52.9%; Pred. No. 1.1e-41;
 Matches 110; Conservative 44; Mismatches 46; Indels 8; Gaps 2;

QY 21 QRPFLIGSGGTASGKSTVCEKIMELLGQNEVEQRQRKVITLSODRFYKVLTAEQAKAL 80
 Db 62 KHPFVIGVCGSASGKTTVAEKIVERLG-----IPWTLISMDSPYKVLPEEIKAAH 114
 QY 81 KGQYNFDPHDPAFNDLMHRTLKNIVEGKTVFPTVDFTVHSRLPETTVYPADVLFEGI 140
 Db 115 ESRYNFDPGNFAPFDLLYEVLKRLREGKSDVPYDFNTHSRDPSNKKMGYADVLFEGI 174
 QY 141 LVFYSQEIRDMFHLRLFVDTSDVRLSRVLRDVR-RGRDLEQLTYTTTVKPAFEFEC 199
 Db 175 LAFHDERIKNLMKMKVFVDTGDLRLARRIVRVDVTRGRDIDGIMEQYFTFVKPAFDKYI 234
 QY 200 LPTKRYADVIFPRGVDNMVAINLIYQHI 227
 Db 235 APCMDSDADLVPRGENDVAIDMIVQNV 262

RESULT 8

Q9U317 PRELIMINARY; PRT; 555 AA.
 AC Q9U317;
 DT 01-MAY-2000 (TREMblrel. 13, Created)
 DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
 DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)
 DE F19B6.1b protein.

GN F19B6.1 OR F19B6.1B.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 CX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Thomas K.;
 RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99069613; PubMed=9851916;
 RA none;
 RT "Genome sequence of the nematode C. elegans: A platform for
 RT investigating biology";
 RL Science 282:2012-2018(1998).
 DR EMBL; Z69635; CA93462.1; -.
 DR HSSP; Q26998; 1BD3.
 DR WormPep; F19B6.1b; CE20700.
 DR InterPro; IPR006082; PRK.
 DR InterPro; IPR000764; Uridine_kin.
 DR Pfam; PF00485; PRK.1.
 DR PRINTS; PR00478; PHRIBLKINASE.
 DR PRINTS; PR00988; URIDINKINASE.
 DR TIGRFAMs; TIGR00235; udk; 1.
 SQ SEQUENCE 555 AA; 62673 MW; D0786FAB98B8CF98 CRC64;

Query Match 38.6%; Score 559; DB 5; Length 555;
 Best Local Similarity 52.9%; Pred. No. 1.3e-41;
 Matches 110; Conservative 44; Mismatches 46; Indels 8; Gaps 2;

QY 21 QRPFLIGSGGTASGKSTVCEKIMELLGQNEVEQRQRKVITLSODRFYKVLTAEQAKAL 80
 Db 102 KHPFVIGVCGSASGKTTVAEKIVERLG-----IPWTLISMDSPYKVLPEEIKAAH 154
 QY 81 KGQYNFDPHDPAFNDLMHRTLKNIVEGKTVFPTVDFTVHSRLPETTVYPADVLFEGI 140
 Db 155 ESRYNFDPGNFAPFDLLYEVLKRLREGKSDVPYDFNTHSRDPSNKKMGYADVLFEGI 214
 QY 141 LVFYSQEIRDMFHLRLFVDTSDVRLSRVLRDVR-RGRDLEQLTYTTTVKPAFEFEC 199
 Db 215 LAFHDERIKNLMKMKVFVDTGDLRLARRIVRVDVTRGRDIDGIMEQYFTFVKPAFDKYI 274
 QY 200 LPTKRYADVIFPRGVDNMVAINLIYQHI 227
 Db 275 APCMDSDADLVPRGENDVAIDMIVQNV 302

RESULT 9

Q8VYB2 PRELIMINARY; PRT; 466 AA.
 AC Q8VYB2;
 DT 01-MAR-2002 (TREMblrel. 20, Created)
 DT 01-MAR-2002 (TREMblrel. 20, Last sequence update)
 DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)
 DE Putative uracil phosphoribosyltransferase.
 GN AtUGS5810.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 CX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Yamada K., Banh J., Chan M.M., Chang C.H., Chang E., Dale J.M.,
 RA Deng J.M., Goldsmith A.D., Lee J.M., Onodera C.S., Quach H.L.,
 RA Tang C., Tortum M., Wu H.C., Yamamura Y., Yu G., Bowser L.,
 RA Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
 RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
 RA Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,
 RA Sakurai T., Satou M., Seki M., Shim P., Southwick A., Shinzaki K.,
 RA Davis R.W., Ecker J.R., Theologis A.;

RT	"Full Length cDNA of gene At1555810 (GI:15222778)";
RL	Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
RL	121
RA	SEQUENCE FROM N.A.
RP	Yamada K., Chan M.M., Chang C.H., Dale J.M., Deng J.M., Hsuan V.M.,
RA	Lee J.M., Quach H.L., Tang C., Tottum M., Wu H.C., Yu G., Bowser L.,
RA	Carninci P., Chen H., Cheuk R., Hayaishizaki Y., Ishida J., Jones T.,
RA	Kamiya A., Karlin-Neumann G., Kawai U., Kim C., Lam B., Lin J.,
RA	Miranda M., Natsuka M., Nguyen M., Palm C.J., Sakurai T., Satou M.,
RA	Seki M., Shim P., Southwick A., Shinozaki K., Davis R.W., Ecker J.R.,
RA	Theologis A.;
RT	"Arabidopsis Open Reading Frame (ORF) Clones";
RL	Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR	EMBL; AY072218; AL50039.1; -
DR	EMBL; AY122946; AA67479.1; -
DR	InterPro; IPR006083; PRK URK.
DR	InterPro; IPR00764; Uridine_kin.
DR	PFam; PF00485; PRK; 1.
DR	PRINTS; PRO0988; URIDINKINASE.
DR	TIGRFAMs; TIGR00235; udk; 1.
DR	Glycoyltransferase; Transferase.
SR	SEQUENCE 466 AA; 52443 MW; 3FA6783CB49727F7 CRC64;
QY	Query Match 37.5%; Score 544; DB 10; Length 466;
Db	Best Local Similarity 43.6%; Pred. No. 2,26-40;
Matches	116; Conservative 52; Mismatches 78; Indels 20; Gaps 8
QY	11 SPAE---ADRRPHRFLIGVSGGASGKSTVCSEKIMELGQNEVEORQRKVVILSDRF 67
Db	29 SNRPBQMAEEHEHGPVIVIGVAGSAASKTYCDIMNOL-----HDQRAVVVNDSF 81
QY	68 YKVLTAQKAKALQGQVNFDPHDAFNDMLMHTLLKNIVEGKTEVPYDPVHSR--LPE 125
Db	82 YHNVEVEIVRV--HDVNFDPHDAFDTQGLSSMKLRKQGVNDIPNDFKSYKNNVPP 139
QY	126 TTVVVPADVLFEGILVFYSQEIQRPMHLRLFVDPDSVRLSRVLRD-YRGRDLEQL 184
Db	140 RR-VNPSDVILLLEGLIFHDPDVRMLMMKIFVDADAVRLARRIKRQVTEGRIATVL 198
QY	185 TQYTFVPAPAEFEFCPTKTKYADVIRPGVDNMVAIVLQVHIDIL-NGDICKNHRGGS 243
Db	199 DQSGFVGPAPBEDFLPTKKTADILIPPGGDNHVAIDLIVGHITKLGQHDCLKLY--P 255
QY	244 NGRSYKRTSEPGDHPGMLTGSKRSH 269
Db	256 NLVYIGSTFGQIRGMHTLLRDSKTTGH 281
RESULT 10	
Q9V810	PRELIMINARY; PRT; 614 AA.
Q9V810	Q9V810; Q9V811;
AC	09V810; Q9V811;
DT	01-MAY-2000 (TrEMBLrel. 13, Created)
DT	01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT	01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE	CG4798 protein.
GN	CG4798.
OS	Drosophila melanogaster (Fruit fly).
OC	Eukaryota, Metazoa, Arthropoda, Hexapoda, Insecta, Pterygota;
OC	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC	Ephydroidea; Drosophilidae; Drosophila.
NCBI_taxid=7227;	
NCBI_taxid=7227;	
SEQUENCE FROM N.A. (LONG AND SHORT ISOFORMS).	
STRAIN=BERKELEY.	
MEDLINE=20196006; PubMed=10731132;	
RA	Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA	Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA	Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA	Brandon R.C., Rogers Y.-H.C., Blasej R.G., Champagne M., Pfeiffer B.D.,
RA	Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA	Abtill J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,

RA	Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA	Beeson K.Y., Benos P.V., Bernan E.P., Bhandari D., Bolshakov S.,
RA	Bortova D., Botchan M.R., Bouck Y., Brokstein P., Brotlier P.,
RA	Butterly K.C., Bussem D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA	Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA	de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA	Dodson K., Dopp L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA	Dutrin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA	Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glaeser K.,
RA	Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA	Harris N.L., Harvey D., Heitman T.J., Hernandez J.R., Houck J.,
RA	Hoslin D., Houston K.A., Howland T.J., Wei M.-H., Ibeagwu C.,
RA	Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison U.A., Ketchum K.A.,
RA	Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulip D., Lai Z.,
RA	Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA	Liu X., Mettel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA	Merkulov G., Milshina N.V., Mobarry C., Morris J., Mosheroff A.,
RA	Mount S.M., Moy M., Murphy B., Murphy L., Muzy D.M., Nelson D.L.,
RA	Nelson D.R., Nelson K.A., Nixon K., Nusserkin D.R., Paclik J.M.,
RA	Palazzolo M., Plittman G.S., Pan S., Pollard J., Putri V., Reese M.G.,
RA	Rahner K., Remington K., Saunders R.D.C., Scheller F., Shen H.,
RA	Shue B.C., Siden-Kiamos I., Simpson M., Skupsky M.P., Smith T.,
RA	Spieler B., Spreading A.C., Stapleton M., Strong R., Sun E.,
RA	Svirskas R., Tecor C., Turner R., Venter E., Wang A.H., Wang X.,
RA	Wang Z.-Y., Wasserman D.A., Weinstock G.W., Weissbach U.,
RA	Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA	Ye Y., Yeh R.-F., Zaveri J.S., Zhao M., Zhang G., Zhao Q., Zheng L.,
RA	Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA	Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT	"The genome sequence of <i>Drosophila melanogaster</i> .";
RL	Science 287:2185-2195(2000).
CC	-I ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A
CC	SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
DR	EMBL: AE003803; AAFF5871.1; -
DR	EMBL: AE003803; AAFF5873.1; -
DR	HSSP: Q26998; 1BD3.
DR	Plybase; FBgn0034213; CG4798.
DR	InterPro; IPR006083; PRK_URK.
DR	InterPro; IPR000764; Uridine_kin.
DR	Pfam; PF00485; PRK; 1.
DR	PRINTS; PR00988; URIDINKINASE.
DR	TIGRFams; TIGR00235; udk; 1.
DR	Alternative splicing; Hypothetical protein.
FT	VARSPLIC 1 207 MISSING (IN SHORT ISOFORM).
FT	SEQUENCE 614 AA; 68704 MW; C81B6B3B9DB392A1 CRC64;
SR	
Query Match	37.4%; Score 542; DB 5; Length 614;
Best Local Similarity	51.5%; Pred. NO. 4.8e-40;
Matches 106; Conservative	42; Mismatches 50; Indels 8; Gaps 3
Oy	23 PELGVSGTASGATVCEKIMELLGQNEVBQRPQRKVILISDPFRVKVLTAEOKAKALKG 82
Db	174 PFLVGICGGSSAGSKTTAAKKIEES--DVW---VTLLMDCPYKLNEKQHQAIIIN 226
Oy	83 QYNPDHPDAPDNMDMHTLTKNIIVEGKTVEPTVDVFTVHSRLPETTVVPADVLFEGLIV 142
Db	227 EYNFDHPADPFIELLDLVLTCLKEGRAKEVAVNPFVTHGRSSQIKTMVGANVIIFEGILT 286
Oy	143 FYSGEIRDMFLRLFLVDTDSVRLSRVLRDV-RRGRLDEQLITQYTTFVKPAFEFCLP 201
Db	287 FHSPEVELTLKMLKFVDDPDIRLARLRARDISQGRDLKVLKOYLMMVKSYSNCYIAP 346
Oy	202 TKKYADVPIIPRGVNMAVINLIYOH 227
Db	347 TMAHDIIIPRGDKNAIHILIVGHV 372
RESULT 11	
ID	OSMOCK4 PRELIMINARY; PRT; 626 AA.
AC	OSMOCK4.
DT	01-OCT-2002 (TREMBLErel. 22, Created)
DT	01-OCT-2002 (TREMBLErel. 22, last sequence update)

DT 01-MAR-2003 (TReMBLrel. 23, last annotation update)
 DE LID03595P.
 GN CG4798.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Peerygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NC NCB1_TaxID=7227;
 RX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkeley;
 RA Stapleton M., Brokstein P., Hong L., Agdayani A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
 RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,
 RA Miranda A., Mungall C.J., Nunco J., Pacle J., Parasas V., Park S.,
 RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
 RA Cejner S.,
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY129436; AAM76178.1; -.
 DR FlyBase; FBgn0034213; CG4798.
 DR InterPro; IPR006083; PRK URK.
 DR InterPro; IPR000764; Uridine_kin.
 DR Pfam; PF00485; PRK.1.
 DR PRINTS; PRO0988; URIDINKINASE.
 DR TIGRFAMs; TIGR00235; udk; 1.
 SQ SEQUENCE 626 AA; 70001 MW; 9D8EC0A7331A6EF3 CRC64;

Query Match 37.4%; Score 542; DB 5; Length 626;
 Best Local Similarity 51.5%; Pred. No. 5e-40; Mismatches 50; Indels 8; Gaps 3;

Matches 106; Conservative 42; Mismatches 50; Indels 8; Gaps 3;
 QY 23 PFLIGSGGTASGKSTVCEKIMELIGONEVEQROKRVILISQDFYKVLTAEQAKALKG 82
 DB 186 PFVIGICGSGASGKTTVAEKIIESE--DVPW---VLLSMDFCYKLINKEQALIN 238
 QY 83 QYNVDHPAFNDLMHRTLKNIVEGKTVETPTVFTVHSRLPETTVYPADVLFEGLIV 142
 DB 239 EYNVDHPAFNFIELLDVLTCLKGRKVEVYVNFVTHGRSQRTWGANVITFEGLIT 298
 QY 143 FYSGEIRDMFHLRFVDTSDVRLSRVLRDV-RRGRDLEILIQYTTVFAPAEFGCLP 201
 DB 299 FHSPEVFLKLMKIFVDTDPDIRLARRLRDISORGRDLKQVLYQVLMVAPSYCNYAP 358
 QY 202 TKKYADVILPRGVDMVAIMLIYOH 227
 DB 359 TMAHADIVPRGDKVAIHILYOHV 384

RESULT 12

Q8N524 PRELIMINARY; PRT; 419 AA.
 AC Q8N524;
 DT 01-OCT-2002 (TReMBLrel. 22, Created)
 DT 01-OCT-2002 (TReMBLrel. 22, last sequence update)
 DT 01-MAR-2003 (TReMBLrel. 23, last annotation update)
 DE Similar to uridine kinase-like 1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NC NCB1_TaxID=9606;
 RX NCB1_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RA Strubeberg R.;
 RL EMBL; BC033078; AAM33078.1; -.
 DR InterPro; IPR006082; PRK.
 DR InterPro; IPR006083; PRK URK.
 DR InterPro; IPR000764; Uridine_kin.
 DR Pfam; PF00485; PRK.1.
 DR PRINTS; PRO0478; PHRIDINKINASE.
 DR PRINTS; PRO0988; URIDINKINASE.
 DR TIGRFAMs; TIGR00235; udk; 1.

KW Kinase.
 SQ SEQUENCE 419 AA; 45972 MW; 1D1F1E75305AD25 CRC64;

Query Match 37.3%; Score 540.5; DB 4; Length 419;
 Best Local Similarity 45.3%; Pred. No. 3.9e-40;
 Matches 110; Conservative 46; Mismatches 60; Indels 27; Gaps 5;

QY 9 CESFAP-----EADRP-----HQRPFILVSGGTASGKSTVCEKIMELIGQ 49
 DB 65 CKSEPPILRTSKRTITVAGRPWYNEHGTOSKAPFALIGGSGASGKTTVARRITEL-- 122
 QY 50 NEVEQROKRVILISQDFYKVLTAEQAKALKCQYNFDPDAPFNDLMHRTLKNIVEGKT 109
 DB 123 -DVPW-----VLLSMDFCYKLTQEQEQAHHNNFNDHPDAPFDLITLTKKQKGS 177
 QY 110 VEVPTVFTVHSRLPETTVYPADVLFEGLIVFYSGEIRDMFHLRFVDTSDVRLSR 169
 DB 178 VKVPIYDFTTHSRKKMKITLYGANVITFEGLIMFARKTLEILDMKIFVDTSDIRLVR 237
 QY 170 VLADV-RRGRDLKQILQYTTFPVKPAFEFECLPTKKYADVILPRGVDMVAIMLIYOH 228
 DB 238 LRRLDISERGRDIEGVLYQVLYQVLYQVLYQVLYQVLYQVLYQVLYQVLYQVLYQ 297
 QY 229 DIL 231
 DB 298 SQL 300

RESULT 13

Q8LD95 PRELIMINARY; PRT; 469 AA.
 AC Q8LD95;
 DT 01-OCT-2002 (TReMBLrel. 22, Created)
 DT 01-OCT-2002 (TReMBLrel. 23, last sequence update)
 DT 01-MAR-2003 (TReMBLrel. 23, last annotation update)
 DE Putative uracil phosphoribosyl transferase.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 NC NCB1_TaxID=3702;
 RX NCB1_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Haas B.J., Volfovsky N., Town C.D., Troupkan M., Alexandrov N.,
 RA Feldmann K.A., Flavell R.B., White O., Salzberg S.L.;
 RT "Full-length messenger RNA sequences greatly improve genome
 RT annotation.";
 RL Genome Biol. 0:0-0(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Brover V., Troupkan M., Alexandrov N., Lu Y.-P., Flavell R.,
 RA Feldmann K.;
 RT "Full-length cDNA from Arabidopsis thaliana.";
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY086133; AAM63338.1; -.
 DR InterPro; IPR006083; PRK URK.
 DR InterPro; IPR000764; Uridine_kin.
 DR Pfam; PF00485; PRK.1.
 DR PRINTS; PRO0988; URIDINKINASE.
 DR TIGRFAMs; TIGR00235; udk; 1.
 KW Transferrase.

Query Match 36.5%; Score 529; DB 10; Length 469;
 Best Local Similarity 48.9%; Pred. No. 4.9e-39;
 Matches 107; Conservative 47; Mismatches 53; Indels 12; Gaps 5;

QY 21 QRPFLIGSGGTASGKSTVCEKIMELIGONEVEQROKRVILISQDFYKVLTAEQAKAL 80
 DB 47 RQFVIVGAGGAASGKTTVCDMLIQQL-----HQRVVLINLBSFYNNLFEELARV- 98
 QY 81 KQYNFDPDAPFNDLMHRTLKNIVEGKTVETPTVFTVHSRLPETTVYPADVLFEGLI 140

```

Db      99 -HEYNFDHPDAPTEHLLSCMEKLROGAQVADIPKDYFKTY-RSSVFRVNPVTVIIIEGI 156
Qy      141 LVFYSOEIDRMFHLRLFVDTDSVRLSRRLRD-VRRGRDLEQILTOYTFVVPAPBEEFC 139
Db      157 LFFHDRVRLKMMKJFVCTDADVRLARRIKRDTVNGRDIQVLDQYSKFVPAFDDFI 216
Qy      200 LPTKRYADVLPFGVNDMMVAINLIVOHIDIL-NGDICK 237
Db      217 LPTKRYADIIIPRGDNNHVAIDLIVGHICTKLGQHDICK 255

RESULT 14
Q9LGF32 PRELIMINARY; PRT; 542 AA.
ID      09LGF32
AC      09LGF32;
DT      01-OCT-2000 (TEMBLrel. 15, Created)
DT      01-OCT-2000 (TEMBLrel. 15, Last sequence update)
DT      01-OCT-2002 (TEMBLrel. 22, Last annotation update)
DE      F14J16.5.
OS      Arabidopsis thaliana (Mouse-ear cress).
OC      Burkaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC      Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC      eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX      NCBI_TaxID=3702;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Shim P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C., Khan S.,
RA      Kim C., Altafi H., Bel O., Chin C., Chou J., Choi E., Conn L.,
RA      Conway A., Gonzales A., Hansen N., Howing B., Koo T., Lam B., Lee J.,
RA      Lenz C., Li J., Liu A., Liu K., Liu S., Mukharasy N., Nguyen M.,
RA      Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaveri A.,
RA      Toriumi M., Vaysberg M., Yu G., Federspiel N.A., Theologis A.,
RA      Ecker J.R.,
RT      "Genomic sequence for Arabidopsis thaliana BAC F14J16 from chromosome
RT      1."
RN      Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
RP      [2]
RA      Ecker J.R.;
RN      Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
RP      [3]
RA      SEQUENCE FROM N.A.
RA      Ecker J.R.;
RN      Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
RP      [4]
RA      SEQUENCE FROM N.A.
RA      Ecker J.R.;
RN      Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
RP      [5]
RA      SEQUENCE FROM N.A.
RA      Ecker J.R.;
RN      Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
RP      [6]
RA      SEQUENCE FROM N.A.
RA      Cheuk R., Shim P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C.,
RA      Khan S., Kim C., Altafi H., Bel O., Chin C., Chou J., Choi E., Conn L.,
RA      Conn L., Conway A., Gonzales A., Hansen N., Howing B., Koo T., Lam B.,
RA      Lee J., Lenz C., Li J., Liu A., Liu K., Liu S., Mukharasy N.,
RA      Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,
RA      Thaveri A., Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N.,
RA      Theologis A., Ecker J.,
RN      Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
RA      EMBL; ACC002304; AAF79310.1; -.
DR      HSSP; Q26996; 1BD3.
DR      InterPro; IPR006083; PRK URK.
DR      InterPro; IPR00764; Uridine_kin.
DR      Pfam; PF00485; PRK.1.
DR      PRINTS; PRO0986; URIDINKINASE..
SQ      SEQUENCE 542 AA; 61022 MW; 35427303EB7F3942 CRC64;

```

```

Qy      11 SPAPB---ADRPDRPEFL-----GVSGGTASGKSTVCEKIMELQNEVEQROKVVILS 63
Db      29 SNRPEQMAEEBEGQFVVGESRGVAGNAGSTTVICDMIMQDL-----HDQRAVVNA 81
Qy      64 QDRFYKVLTAEOKAKALKQVNFDPDAPDNMLMRLTNIVEGKTEVEPTVDFVHSR- 122
Db      82 QDSFHNNVNEVELVR--HDVNFDPDAPDTQLLSMCKLKGQAVDLPNVDKSKYKN 139
Qy      123 -LPETTVYPADVLPFEGILVFSQETIRDMFHLRFVDT----- 160
Db      140 VFEPFR-VNPSDVIILEGILIFHDPVRDMMKIFVDAGLSHTKPEVNTYVKSVAVMR 198
Qy      161 -----DSVRLSRVRD-VRRGRDLEQILTOYTFVVPAPBEEFCPTKRYADVLP 211
Db      199 CTCICTHEDADVRLARRIKRDTVNGRDIQVLDQYSKFVPAFDDFIPTKRYADIIIP 258
Qy      212 RGVNDMMVAINLIVOHIDIL-NGDICKMHRGSGNGSRKRTSEPDHGMILTSGRSH 269
Db      259 RGDNNHVAIDLIVGHHTLGLGQHDICKIY---PNLVIOSTPQIRGMHLLIDSKTTKH 314

RESULT 15
Q9LFF22 PRELIMINARY; PRT; 1060 AA.
ID      Q9LFF22
AC      Q9LFF22;
DT      01-OCT-2000 (TEMBLrel. 15, Created)
DT      01-OCT-2000 (TEMBLrel. 15, Last sequence update)
DT      01-OCT-2002 (TEMBLrel. 22, Last annotation update)
DE      F20N2.19.
OS      Arabidopsis thaliana (Mouse-ear cress).
OC      Burkaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC      Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC      eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX      NCBI_TaxID=3702;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Khan S., Brooks S., Buehler E., Chao Q., Johnson-Hopson C., Kim C.,
RA      Shim P., Altafi H., Bel O., Chin C., Chou J., Choi E., Conn L.,
RA      Conway A., Gonzales A., Hansen N., Howing B., Koo T., Lam B., Lee J.,
RA      Lenz C., Li J., Liu A., Liu K., Liu S., Mukharasy N., Nguyen M.,
RA      Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaveri A.,
RA      Toriumi M., Vaysberg M., Yu G., Federspiel N.A., Theologis A.,
RA      Ecker J.R.,
RT      "Genomic sequence for Arabidopsis thaliana BAC F20N2 from chromosome
RT      1."
RN      Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
RP      [2]
RA      SEQUENCE FROM N.A.
RA      Ecker J.R.;
RN      Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
RP      [3]
RA      SEQUENCE FROM N.A.
RA      Ecker J.R.;
RN      Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
RP      [4]
RA      SEQUENCE FROM N.A.
RA      Walker P., Buehler E.E., Dunn P.P., Feng J.J., Kim C.C., Li Y.Y.,
RA      Walker M.M., Altafi H.H., Araujo R.R., Conn L.L., Conway A.A.B.,
RA      Gonzales A.A., Hansen N.N.F., Huizar L.L., Kremetskaya I.I.,
RA      Lenz C.C., Li J.J., Liu S.S., Lueros S.S., Rowley D.D., Schwartz J.J.,
RA      Toriumi M.M., Vysotskaia V.V., Yu G.G., Davis R.R.W.,
RA      Federspiel N.N.A., Theologis A.A., Ecker J.J.R.,
RN      Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
RP      [5]
RA      SEQUENCE FROM N.A.
RA      Ecker J.R.;
RN      Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
RP      [6]
RA      SEQUENCE FROM N.A.
RA      Cheuk R., Shim P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C.,
RA      Khan S., Kim C., Altafi H., Bel O., Chin C., Chou J., Choi E.,
RA      Conn L., Conway A., Gonzales A., Hansen N., Howing B., Koo T., Lam B.,

```

```

Query Match      35.0%; Score 507.5; DB 10; Length 542;
Best Local Similarity 38.8%; Pred. No. 5,1e-37;
Matches 116; Conservative 52; Mismatches 78; Indels 53; Gaps 10;

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2003, 07:59:59 ; Search time 12 Seconds

(without alignments)
1085.533 million cell updates/sec

Title: US-09-896-522-2

Perfect score: 1450
Sequence: 1 MASAGSDSCSPAPEADRP.....HPGWLTSGRKSHLESSSRPH 277

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1450	100.0	277	1 UCK1_HUMAN	Q9HA47 homo sapien
2	1325	91.4	277	1 UCK1_MOUSE	P52623 homo musculu
3	934	64.4	261	1 UCK2_HUMAN	Q9BZX2 homo sapien
4	932	64.3	261	1 UCK2_MOUSE	Q9BPM9 mus musculu
5	738	50.9	260	1 UCK_DROME	Q9VCS9 drosophila
6	545.5	37.6	548	1 URL1_HUMAN	Q9NW35 homo sapien
7	539.5	37.2	548	1 URL1_MOUSE	Q9IY13 mus musculu
8	512.5	35.3	229	1 UCK_CAEEL	Q17413 caenorhabdi
9	473	32.6	211	1 URK_BACHD	Q9K488 bacillus ha
10	468	32.3	208	1 URK_CLOPE	Q8XJ16 clostridium
11	454	31.3	206	1 URK_LACLA	Q9CF21 lactococcus
12	447	30.8	207	1 URK_STAM	Q9Y1T8 staphylococ
13	439.5	30.3	208	1 URK_STRP8	Q9P018 streptococc
14	439.5	30.3	208	1 URK_STRP	Q9P270 streptococc
15	438	30.2	212	1 URK_STRP	Q9J7J7 streptococc
16	437	30.1	501	1 URK1_YEAST	P27515 saccharomyc
17	430	29.7	211	1 URK_BACSU	Q32033 bacillus su
18	417	28.8	225	1 URK_HALTU	Q9HGS9 halobacteri
19	412	28.4	222	1 URK_CHLNP	Q9Z7H0 chlamydia p
20	410.5	28.3	205	1 URK_BORBU	Q59190 borrelia bu
21	408	28.1	210	1 URK_DEIRA	Q9RX55 diinococcu
22	406	28.0	213	1 URK_VIBCH	Q9K677 vibrio chol
23	398	27.4	209	1 URK_LISIN	Q9ZB16 listeria in
24	392	27.0	209	1 URK_LISMO	Q9Y777 listeria in
25	384	26.5	213	1 URK_HAELN	P44533 haemophilus
26	377.5	26.0	213	1 URK_SALTY	Q8XEY2 salmonella
27	375.5	25.9	213	1 URK_ECOLI	P31218 escherichia
28	371.5	25.6	213	1 URK_MYCPN	P75217 mycoplasma
29	363.5	25.1	213	1 URK_MYCCE	P47622 mycoplasma
30	362.5	25.0	213	1 URK_YERPE	Q8ZTF5 yersinia pe
31	362	25.0	216	1 URK_PASMU	Q9ZCM5 pasteurella
32	328.5	22.7	207	1 URK_UREPA	Q9PGF9 ureaplasma
33	217	15.0	404	1 KPPR_WHEAT	P26302 triticum ae

34	205.5	14.2	395	1 KPPR_ARATH	P25697 arabidopsis
35	203	14.0	332	1 KPPR_SYNY3	P37101 synchocyst
36	200	13.8	397	1 KPPR_MESCR	P27774 mesembryant
37	195.5	13.5	402	1 KPPR_SPIOL	P09559 spiracila ol
38	183	12.6	322	1 COAA_AGRTS	Q8UJ92 agrobacteri
39	183	12.6	375	1 KPPR_CHLRE	P19824 chlamydomon
40	180.5	12.4	319	1 COAA_BACSU	P54536 bacillus su
41	178	12.3	322	1 COAA_BRUME	Q8Y639 brucella me
42	175.5	12.1	306	1 COAA_LISMO	Q8Y810 listeria mo
43	171	11.8	316	1 COAA_PASMU	P57967 pasteurella
44	170.5	11.6	306	1 COAA_LISIN	Q9ZB16 listeria in
45	166	11.4	331	1 COAA_RHIME	Q9ZBT5 rhizobium m

ALIGNMENTS

RESULT 1
ID UCK1_HUMAN STANDARD; PRT; 277 AA.
AC Q9HA47;
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Uridine-cytidine kinase 1 (EC 2.7.1.48) (UCK 1) (Uridine monophosphokinase 1) (Cytidine monophosphokinase 1).
GN UCK1.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
RX MEDLINE=21203813; PubMed=11306702;
RA Van Rompay A.R., Norda A., Linden K., Johansson M., Karlsson A.;
RT "Phosphorylation of uridine and cytidine nucleoside analogs by two human uridine-cytidine kinases.";
RL Mol. Pharmacol. 59:1181-1186(2001).
RN [2]
RP SEQUENCE FROM N.A.
RA Ho Y.S., Johnson R.K.;
RT "Human uridine kinase from prostate cancer cell line (LNCap)." ;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA TISSUE=Mammary gland;
RC Ilogai T., Ota T., Hayaishi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
RA Nagatsuna M., Hosoi T., Kaku Y., Kodaira H., Kondo H.,
RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
RA Yamamoto J., Wakamatsu A., Nakamura K., Ishi S., Kawai Y., Saito K.,
RA Niimura K., Iwayanagi T.;
RT "NEDD human cDNA sequencing project." ;
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Xin Y.R., Yu L., Zhao S.Y.;
RT "Cloning of a new human cDNA similar to Mus musculus uridine kinase mRNA." ;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Phosphorylates uridine and cytidine to uridine monophosphate and cytidine monophosphate. Does not phosphorylate deoxyribonucleosides or purine ribonucleosides. Can use ATP or GTP as a phosphate donor. Can also phosphorylate cytidine and uridine nucleoside analogs such as 6-azauridine, 5-fluorouridine, 4-thiouridine, 5-bromouridine, N(4)-acetylcytidine, N(4)-benzoylcytidine, 5-fluorocytidine, 2-thiocytydine, 5-methylcytidine, and N(4)-antioylcytidine.
CC -!- CATALYTIC ACTIVITY: ATP + uridine = ADP + UMP.
CC -!- CATALYTIC ACTIVITY: ATP + cytidine = ADP + CMP.
CC -!- PATHWAY: Pyrimidine salvage pathway.
CC -!- TISSUE SPECIFICITY: Ubiquitous.

```

CC -1- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF237290; AAK28324.1; -
DR EMBL; AF254133; AAK49122.1; -
DR EMBL; AK022317; BAB14010.1; -
DR EMBL; AF125106; AAL75943.1; -
DR InterPro; IPR006083; PRK-URK.
DR InterPro; IPR000764; Uridine_kin.
DR Pfam; PF00485; PRK; 1.
DR PRINTS; PR00988; URIDINKINASE.
DR TIGRFAMs; TIGR00235; udk; 1.
KM Transferase; Kinase; ATP-binding.
FT NP_BIND 30 37 ATP (POTENTIAL).
FT CONFLICT 8 17 DCESPAPBAD -> GARAPAGAN (IN REF. 4).
FT CONFLICT 56 57 OR -> HG (IN REF. 4).
FT CONFLICT 247 247 S -> T (IN REF. 4).
SQ SEQUENCE 277 AA; 31434 MW; AFD9ED92780CD502 CRC64;

Query Match 100.0%; Score 1450; DB 1; Length 277;
Best Local Similarity 100.0%; Pred. No. 5,7e-113;
Matches 277; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASAGDECCSPAPADRPORPFLIGVSGTASGKTVCEKIMELLGONEVEORORRV 60
DB 1 MASAGDECCSPAPADRPORPFLIGVSGTASGKTVCEKIMELLGONEVEORORRV 60
QY 61 ILSODRFYKVLTAEOAKALKGOYNFDPADFPNDLMHRTLKNIVGKTVETVYDFVTH 120
DB 61 ILSODRFYKVLTAEOAKALKGOYNFDPADFPNDLMHRTLKNIVGKTVETVYDFVTH 120
QY 121 SRLPFTTVVPADVVLFEGLIVFYSGEIRDMFHLRLFVDTSDVRLSRVLRDVRGRDL 180
DB 121 SRLPFTTVVPADVVLFEGLIVFYSGEIRDMFHLRLFVDTSDVRLSRVLRDVRGRDL 180
QY 121 SRLPFTTVVPADVVLFEGLIVFYSGEIRDMFHLRLFVDTSDVRLSRVLRDVRGRDL 180
DB 121 SRLPFTTVVPADVVLFEGLIVFYSGEIRDMFHLRLFVDTSDVRLSRVLRDVRGRDL 180
QY 181 EQLITQYTFYKPAFEEFCPLTKKYADVIIPRGVNMVAIILVQHIQDILNGDICKMHR 240
DB 181 EQLITQYTFYKPAFEEFCPLTKKYADVIIPRGVNMVAIILVQHIQDILNGDICKMHR 240
QY 241 GGSNGRSYKRTFSEPGDHGMILTSKRSKSHLESSSRPH 277
DB 241 GGSNGRSYKRTFSEPGDHGMILTSKRSKSHLESSSRPH 277

RESULT 2
UCCK1 MOUSE STANDARD; PRT; 277 AA.
ID UCCK1 MOUSE
AC P52623;
DT 01-OCT-1996 (Rel. 34, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Uridine-cytidine kinase 1 (EC 2.7.1.48) (UCCK 1) (Uridine
DE monophosphokinase 1) (Cytidine monophosphokinase 1).
GN UCCK1 OR UMPK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Salivary gland;
RX MEDLINE=22388257; PubMed=12477932;
RA Klausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shemmen C.M., Schuler G.D.,
RA Aleschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

```

```

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Pange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McKernan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Hailton E., Ketterman M., Madan A., Rodrigues S., Sanchez A.,
RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences."
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE OF 18-277 FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=97108719; PubMed=8951040;
RA Kopp P.A., Traut T.W.;
RT "Cloning and expression of a cDNA encoding uridine kinase from mouse
RT brain."
RL Arch. Biochem. Biophys. 336:105-112(1996).
CC -1- FUNCTION: Phosphorylates uridine and cytidine to uridine
CC monophosphate and cytidine monophosphate. Does not phosphorylate
CC deoxyribonucleosides or purine ribonucleosides. Can use ATP or GTP
CC as a phosphate donor. Can also phosphorylate cytidine and uridine
CC nucleoside analogs such as 6-azauridine, 5-fluorouridine, 4-
CC thioridine, 5-bromouridine, N(4)-acetylcytidine, N(4)-
CC benzoylcytidine, 5-fluorocytidine, 2-thiocytidine, 5-
CC methylcytidine, and N(4)-anisoylcytidine (By similarity).
CC -1- CATALYTIC ACTIVITY: ATP + uridine = ADP + UMP.
CC -1- CATALYTIC ACTIVITY: ATP + cytidine = ADP + CMP.
CC -1- PATHWAY: Pyrimidine salvage pathway.
CC -1- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; BC025146; AAK25146.1; -
DR EMBL; L31783; AAB50568.1; -
DR MGD; MGI:98904; Umpk.
DR InterPro; IPR006083; PRK URK.
DR InterPro; IPR000764; Uridine_kin.
DR Pfam; PF00485; PRK; 1.
DR PRINTS; PR00988; URIDINKINASE.
DR TIGRFAMs; TIGR00235; udk; 1.
KM Transferase; Kinase; ATP-binding.
FT NP_BIND 30 37 ATP (POTENTIAL).
SQ SEQUENCE 277 AA; 31068 MW; 3EBB3C4187FAEB4A CRC64;

Query Match 91.4%; Score 1325; DB 1; Length 277;
Best Local Similarity 91.7%; Pred. No. 1,3e-102;
Matches 254; Conservative 11; Mismatches 12; Indels 0; Gaps 0;

QY 1 MASAGDECCSPAPADRPORPFLIGVSGTASGKTVCEKIMELLGONEVEORORRV 60
DB 1 MASAGDECCSPAPADRPORPFLIGVSGTASGKTVCEKIMELLGONEVEORORRV 60
QY 61 ILSODRFYKVLTAEOAKALKGOYNFDPADFPNDLMHRTLKNIVGKTVETVYDFVTH 120
DB 61 ILSODRFYKVLTAEOAKALKGOYNFDPADFPNDLMHRTLKNIVGKTVETVYDFVTH 120
QY 121 SRLPFTTVVPADVVLFEGLIVFYSGEIRDMFHLRLFVDTSDVRLSRVLRDVRGRDL 180
DB 121 SRLPFTTVVPADVVLFEGLIVFYSGEIRDMFHLRLFVDTSDVRLSRVLRDVRGRDL 180
QY 121 SRLPFTTVVPADVVLFEGLIVFYSGEIRDMFHLRLFVDTSDVRLSRVLRDVRGRDL 180
DB 121 SRLPFTTVVPADVVLFEGLIVFYSGEIRDMFHLRLFVDTSDVRLSRVLRDVRGRDL 180

```

QY 181 EQLITQYTFEKAFAEEFCLEPTKKYADVILIPRGVNNVAINLIVQHIQDILNGDICKMHR 240
 DB 181 EQLITQYTAIVKAFEEFCLEPTKKYADVILIPRGVNNVAINLIVQHIQDILNGDICKMHR 240
 QY 241 GGSNGRSYKRTFSEBPDHPCMLTSGKRSKSHLESSSRPH 277
 DB 241 GGPGRNHRKRTFPEFGDHPGLATGKRSHLESSSRPH 277

RESULT 3
 UCK2_HUMAN STANDARD; PRT; 261 AA.
 ID UCK2_HUMAN
 AC O99BZK2; O99BZK5;
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DE 28-FEB-2003 (Rel. 41, Last sequence update)
 DE Uridine-cytidine kinase 2 (EC 2.7.1.48) (UCK 2) (Uridine monophosphokinase 2) (Cytidine monophosphokinase 2).
 GN UCK2 OR UMPK.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NCBI_TaxID=9606;
 [1]
 RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
 RX MEDLINE=21203813; PubMed=11306702;
 RA Van Rompay A.R., Norda A., Linden K., Johansson M., Karlsson A.;
 RT "Phosphorylation of uridine and cytidine nucleoside analogs by two human uridine-cytidine kinases."
 RL Mol. Pharmacol. 59:1181-1186(2001).
 [2]
 RP SEQUENCE OF 15-261 FROM N.A.
 RC TISSUE=Fibrosarcoma;
 RX MEDLINE=21385121; PubMed=11494055;
 RA Koizumi K., Shiamoto Y., Azuma A., Wataya Y., Matsuda A., Sasaki T., Fukushima M.;
 RT "Cloning and expression of uridine/cytidine kinase cDNA from human fibrosarcoma cells."
 RL Int. J. Mol. Med. 8:273-278(2001).
 CC -1- FUNCTION: Phosphorylates uridine and cytidine to uridine monophosphate and cytidine monophosphate. Does not phosphorylate deoxyribonucleosides or purine ribonucleosides. Can use ATP or GTP as a phosphate donor. Can also phosphorylate cytidine and uridine nucleoside analogs such as 6-azauridine, 5-fluorouridine, 4-thiouridine, 5-bromouridine, N(4)-acetylcytidine, N(4)-methylcytidine, and N(4)-anisoylcytidine.
 CC -1- CATALYTIC ACTIVITY: ATP + uridine = ADP + UMP.
 CC -1- CATALYTIC ACTIVITY: ATP + cytidine = ADP + CMP.
 CC -1- PATHWAY: Pyrimidine salvage pathway.
 CC -1- TISSUE SPECIFICITY: Expressed in placenta.
 CC -1- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
 CC
 CC EMBL: AF236637; AAK14053.1; -
 CC EMBL: AF062451; BAB5162.1; -
 CC Genbank: H08C12362; UMPK.
 CC InterPro: IPR006082; PRK.
 CC InterPro: IPR006083; PRK_URK.
 CC InterPro: IPR00764; Uridine_kin.
 CC Pfam: PF00485; PRK_1.
 CC PRINTS: PR00478; PHRIUKINASE.
 CC PRINTS: PR00988; URIDINKINASE.
 CC TIGRFAMs: TIGR00235; uck; 1.
 CC Transferase: Kinase; ATP-binding.
 CC NP_BIND 27 ATP (POTENTIAL).

SQ SEQUENCE 261 AA; 29299 MW; 71791346F091EBFD CRC64;
 Query Match 64.4%; Score 934; DB 1; Length 261;
 Best Local Similarity 69.5%; Pred. No. 2, 9e-70;
 Matches 189; Conservative 27; Mismatches 40; Indels 16; Gaps 4;

QY 8 DCESPAEADRPPI-QRPFLIGVSGTASGKSTVCEKIMELLGONEVEOROKVILISQDR 66
 DB 4 DSEQTLQNNQPPRGSGPFLIGVSGTASGKSSVCATIVQLGNEVDYRKQVILISQDS 63
 QY 67 FYVLTAEQAKAKLKOQYFNPDHDPNDLMHRTLNIVEGKTEVPYDVTSHSLPET 126
 DB 64 FYVLTSEQAKAKLKOQFNPDHDPNDLMHRTLNIVEGKTEVPYDVTSHSLPET 123
 QY 127 TVVYPPADVLLFEGILVFGQELRDMTHLLFVDTSDVLSRRVADV-RRGDLEQIIL 185
 DB 124 VVYPPADVLLFEGILVFGQELRDMTHLLFVDTSDVLSRRVADV-RRGDLEQIIL 183
 QY 186 QYTFVKAFAEEFCLEPTKKYADVILIPRGVNNVAINLIVQHIQDILNGDICKMHRGSG 245
 DB 184 QYTFVKAFAEEFCLEPTKKYADVILIPRGVNNVAINLIVQHIQDILNGDICKMHRGSG 242
 QY 246 RSYKRTFSEBPDHPCMLTSGKRSKSHLESSSRPH 277
 DB 233 PSKRQT---NGCLNGYTPSKRQASSESSSRPH 261

RESULT 4
 UCK2_MOUSE STANDARD; PRT; 261 AA.
 ID UCK2_MOUSE
 AC O99PM9;
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DE 28-FEB-2003 (Rel. 41, Last sequence update)
 DE Uridine-cytidine kinase 2 (EC 2.7.1.48) (UCK 2) (Uridine monophosphokinase 2) (Cytidine monophosphokinase 2).
 GN UCK2 OR UMPK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21203813; PubMed=11306702;
 RA Van Rompay A.R., Norda A., Linden K., Johansson M., Karlsson A.;
 RT "Phosphorylation of uridine and cytidine nucleoside analogs by two human uridine-cytidine kinases."
 RL Mol. Pharmacol. 59:1181-1186(2001).
 CC -1- FUNCTION: Phosphorylates uridine and cytidine to uridine monophosphate and cytidine monophosphate. Does not phosphorylate deoxyribonucleosides or purine ribonucleosides. Can use ATP or GTP as a phosphate donor. Can also phosphorylate cytidine and uridine nucleoside analogs such as 6-azauridine, 5-fluorouridine, 4-thiouridine, 5-bromouridine, N(4)-acetylcytidine, N(4)-methylcytidine, and N(4)-anisoylcytidine (by similarity).
 CC -1- CATALYTIC ACTIVITY: ATP + uridine = ADP + UMP.
 CC -1- CATALYTIC ACTIVITY: ATP + cytidine = ADP + CMP.
 CC -1- PATHWAY: Pyrimidine salvage pathway.
 CC -1- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
 CC
 CC EMBL: AF236636; AAK14052.1; -
 CC MGD: MGI:1931744; Uck2.
 CC InterPro: IPR006082; PRK.
 CC InterPro: IPR006083; PRK_URK.

DR InterPro: IPR000764; Uridine_kin.
 DR Pfam: PF00485; PRK; 1.
 DR PRINTS: PR00478; PHRIKINASE.
 DR PRINTS: PR00988; URIDINKINASE.
 DR TIGRFAMs: TIGR00235; uck; 1.
 DR Transferrase; Kinase; ATP-binding.
 FT NP BIND 27 ATP (POTENTIAL).
 SQ SEQUENCE 261 AA; 29404 MW; 7800AA3CF5C8153 CRC64;
 Query Match 64.3%; Score 932; DB 1; Length 261;
 Best Local Similarity 68.4%; Pred. No. 4.3e-70;
 Matches 186; Conservative 27; Mismatches 43; Indels 16; Gaps 4;
 QY 8 DCEPAPAEADPH-QRPFLIVSGSGTASGKSTVEKIMELQNEVEQKQKVVITLSDR 66
 DB 4 DSEQTLNQHQPNSEPEPLIVSGSGTASGKSSVCAKIVQLQNGNEVDHQVILLSDS 63
 QY 67 FYKVLTAEQKALKGQYFDPHDPADFDMLHRTLKNIVEGKVEPPTVDFVTHSLPET 126
 DB 64 FYRLTSEQKALKGQYFDPHDPADFDMLHRTLKNIVEGKVEPPTVDFVTHSLPET 123
 QY 127 TVVYPADVLEFEGILVFYSQIRDMFHLRLFDVTDSDVRLSRVLRDV-RRGRDLEQILT 185
 DB 124 VTIVPADVLEFEGILVFYSQIRDMFHLRLFDVTDSDVRLSRVLRDV-RRGRDLEQILT 183
 QY 186 QYTFVPAFAFEFCLPTKRYADVILPRGVDMVAINLIVQHIDILNGDICKHRRGSNG 245
 DB 184 QYTFVPAFAFEFCLPTKRYADVILPRGADVLVAINLIVQHIDILNGLSKRTQNG--- 240
 QY 246 RSYKRTSEPGDHGMLTSGKSHLESSRPH 277
 DB 241 --YLN-----GYTPSRKQASSESSSRPH 261
 RESULT 5
 UCK_DROME STANDARD; PRT; 260 AA.
 AC Q9Y299;
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Probable uridine-cytidine kinase (EC 2.7.1.48) (UCK) (Uridine
 DE monophosphokinase) (Cytidine monophosphokinase).
 GN CG6364.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pezomyzeta;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champs C.M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abrial J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borokova D., Botchan D.A., Bouck J., Brokstein P., Brotlier P.,
 RA Buttle K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA de Paolis B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrieri S., Fleischmann W.,
 RA Foster J.C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hotlin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jaitai M., Kalush F., Karpen G.H., Ke Z., Kemison J.A., Ketchum K.A.,

RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lesko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshirei A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusser D.R., Pacle J.M.,
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Slier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Stykxas R., Teator C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wastaman D.A., Weinstein G.M., Weisenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 CC -1- CATALYTIC ACTIVITY: ATP + uridine = ADP + UMP.
 CC -1- CATALYTIC ACTIVITY: ATP + cytidine = ADP + CMP.
 CC -1- PATHWAY: Pyrimidine salvage pathway.
 CC -1- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: AB003747; AA056274.1; -.
 DR Flybase: FBgn0039179; CG6364.
 DR InterPro: IPR000764; Uridine_kin.
 DR InterPro: IPR000764; Uridine_kin.
 DR Pfam: PF00485; PRK; 1.
 DR PRINTS: PR00988; URIDINKINASE.
 KW Hypothetical protein; Transferase; Kinase; ATP-binding.
 FT NP BIND 34 ATP (POTENTIAL).
 SQ SEQUENCE 260 AA; 29213 MW; 2EE57423704925B6 CRC64;
 Query Match 50.9%; Score 738; DB 1; Length 260;
 Best Local Similarity 67.8%; Pred. No. 5e-54;
 Matches 143; Conservative 32; Mismatches 34; Indels 2; Gaps 2;
 QY 23 PFLIGVSGTASGKSTVCEKIMELQNEVEQKQKVVITLSDRPFYKVLTAEQKALKG 82
 DB 27 PFLIGVAGGTASGKSTVCKKIMEQLQAEWDHTQROGVVSIQDSFYRELTPAKKAKAOKG 86
 QY 83 QYNEPDHDPADFDNDMLHRTLKNIVEGKVEPPTVDFVTHSL-RLPETTVVYPADVLEFEGILT 141
 DB 87 LFNFDHDPADFAEELMTSLQNLKLGKVELPSPDYRTNSLDPPNVLVITPADVLEFEGILT 146
 QY 142 VFVSOEIRDMFHLRLFDVTDSDVRLSRVLRDV-RRGRDLEQILTQYTFVPAFAFEFCL 200
 DB 147 VFVFPKIRLRFHKLRLFDVTDSDVRLSRVLRDV-RRGRDLEQILTQYTFVPAFAFEFCL 206
 QY 201 PTKKFAVDVILPRGADVLVAINLIVQHIDIL 231
 DB 207 PTKKFAVDVILPRGADVLVAINLIVQHIDIL 237
 RESULT 6
 URL1_HUMAN STANDARD; PRT; 548 AA.
 AC Q9NMZ5; Q9H322;
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Uridine kinase-like 1.
 GN URKL1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OX NCBI_TaxID=9606;
 RN
 RP SEQUENCE FROM N.A.
 RA Watanabe K., Kamagai A., Itakura S., Yamazaki M., Tashiro H., Ota T.,
 RA Suzuki Y., Obayashi M., Nishi T., Shibahara T., Tanaka T.,
 RA Nakamura Y., Iwagaki T., Sugano S.;
 RT "MEDO human cDNA sequencing project";
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 RN
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21638749; PubMed=11780052;
 RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
 RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Baggall C.L.,
 RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
 RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
 RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
 RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
 RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,
 RA Coulson A., Cowell G.J., Deadman R., Dhami P.D., Dunn M.,
 RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
 RA Grafham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
 RA Hammond S., Harley J.B., Heath P.D., Ho S., Holden J.L., Howden P.J.,
 RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
 RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
 RA Leharshah M.H., Leverhulme M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
 RA Marsh V.L., Martin S.L., McComachie L.J., McElroy K., McMurtry A.A.,
 RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
 RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
 RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsey H.,
 RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showstren R., Sims S.,
 RA Sruce C.D., Smith M.L., Soderlund C., Steward C.A., Sulten J.E.,
 RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
 RA Tracey A., Tromane A.C., Vaubin M., Wall M., Wallis J.M., Williams S.A.,
 RA Whitehead S.L., Whitlaker P., Willey D.L., Williams L., Williams S.A.,
 RA Wilmshurst L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
 RA Rogers J.;
 RT "The DNA sequence and comparative analysis of human chromosome 20";
 RL Nature 414:865-871(2001).
 CC
 CC -1- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: AK00524; BAB91330.1; -;
 DR EMBL: AL118506; CAC15497.1; -;
 DR HSSP: Q26998; 1BD3.
 DR GeneW: HGNC:15938; URKUL1.
 DR InterPro: IPR006082; PRK.
 DR InterPro: IPR006083; PRK URK.
 DR InterPro: IPR000764; Uridine_kin.
 DR Pfam: PF00485; PRK.1.
 DR PRINTS: PR00478; PHRIDKINASE.
 DR PRINTS: PR00988; URIDINKINASE.
 DR TIGRPFAMs: TIGR00235; udk; 1.
 KM Transferrase; Kinase; ATP-binding.
 FT NP BIND 105 112 ATP (POTENTIAL).
 FT CONFLICT 219 228 MISSING (IN REF. 2).
 FT CONFLICT 290 290 N -> D (IN REF. 2).
 SQ SEQUENCE 548 AA; 61139 MW; 0CD03697E02FE7DB CRC64;
 Query Match 37.64; Score 545.5; DB 1; Length 548;
 Best Local Similarity 45.74; Pred. No. 1.1e-37;
 Matches 111; Conservative 45; Mismatches 60; Indels 27; Gaps 5;
 QY 9 CESAPAP-----EADRP-----HOPFPLIGVGSGTASGKSYCEKIMELGQ 49
 DB 65 CKSEPLFLRTSKRTIYAGRPVPMVNEHGTSKEAFALGDSGSGSKTIVARIIEAL-- 122

QY 50 NEVEQRQRKRVILISODRFYKVLTAEOKAKALGQVNFHDPAFPNDLMARLTKNIEGKT 109
 DB 123 -DVPW---VLLMSDFYKVLLEBOQEOAHNNRFDPDAPFDPLISTLKAKQGS 177
 QY 110 VEPFYDFYTHRLPPTTVVYPADVVLFEGLIVFYSOEIRDFHRLFPVDTSDVRLSPR 169
 DB 178 VNPFIYDFTHRKQKMWLTGNAVIFEGINAFADKLTLELDKMFVDTSDIRLVR 237
 QY 170 VLRDV-RRGRLDEQILTYOTTTVKPAFESECLPTKXADVILPRGVNVAINLYQHIQ 228
 DB 238 LRRDISERGRDIEGVIKQYKRFKESFDQYIQPTWLADIVVPRGSGNTVAINTLVQVNH 297
 QY 229 DIL 231
 DB 298 SQL 300
 RESULT 7
 URL_MOUSE STANDARD; PRT; 548 AA.
 ID URL_MOUSE
 AC 091Y13;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Uridine kinase-like 1.
 GN URKUL1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klauener R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Dlatchenko L., Matovina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton W., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosack S.A., McMan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Wuzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Maitra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC
 CC -1- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: BC016535; AAH16535.1; -;
 DR MGD: MGI:1915806; 1110007H10R1K.
 DR InterPro: IPR006082; PRK.
 DR InterPro: IPR006083; PRK URK.
 DR InterPro: IPR000764; Uridine_kin.
 DR Pfam: PF00485; PRK.1.
 DR PRINTS: PR00478; PHRIDKINASE.
 DR PRINTS: PR00988; URIDINKINASE.
 DR TIGRPFAMs: TIGR00235; udk; 1.

KW Transferrase; Kinase; ATP-binding.
 FT NP BIND 105 112 ATP (POTENTIAL).
 SQ SEQUENCE 548 AA; 60841 MW; 9945CA2359BD525 CRC64;

Query Match 37.2%; Score 539.5; DB 1; Length 548;
 Best Local Similarity 45.3%; Pred. No. 3.6e-37;
 Matches 110; Conservative 46; Mismatches 60; Indels 27; Gaps 5;

QY 9 CESRAP-----EADRP-----HQRPFLIGSGSTVCEKIMELLGQ 49
 DB 65 CKSEPPRLRTSKRTITYTAGRPWVNEHGTSKEAFALIGGSGASGKTVAEMTEAL-- 122
 QY 50 NEVEQQRGRKVVILSDREFYKVLTAEOKAKALGQGNFDPDPAPNDMLHRTLKNIIVEGKT 109
 DB 123 -DVEW----VVLMSDSFYKVLTLQCGQQAACNNFNFPDPAPFDLITSLKLGKGRS 177
 QY 110 VEPYTFDFTHSRLEPPTTVVPADVLFEGILVFSQEIQDMFHLRFVDTSDVRLSRR 169
 DB 178 VQVPIYDFTTSRKMDKMTLYGANVILFEGIMAFADKLTLELLDMKIFVDTSDIRLVRR 237
 QY 170 VLARV-RRGRDLEQLIQYTFEVKPAPEEFCLEPTKKYADVILPRGVNMAINLVQHIO 228
 DB 238 LRRDISERGRDIEGVIKQYKFKVPAFDQYIQTMRADIVPRSGNTVAIDLIVQHVH 297
 QY 229 DIL 231
 DB 298 SQL 300

RESULT 8
 ID UCK CAEEL STANDARD; PRT; 229 AA.
 UCK CAEEL

AC 017413;
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Probable uridine-cytidine kinase (EC 2.7.1.48) (UCK) (Uridine monophosphokinase) (Cytidine monophosphokinase).
 GN B0001.4.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea; Rhabditidae; Pelodierinae; Caenorhabditis.
 OC NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol NZ;
 RA Sims M.;
 RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: ATP + uridine = ADP + UMP.
 CC -1- CATALYTIC ACTIVITY: ATP + cytidine = ADP + CMP.
 CC -1- PATHWAY: Pyrimidine salvage pathway.
 CC -1- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; Z69634; CA93453.1; -.
 DR PIR; T18629; T18629.
 DR WormPep; B0001.4; CE05142.
 DR InterPro; IPR006082; PRK.
 DR InterPro; IPR006083; PRK.
 DR InterPro; IPR00764; Uridine_kin.
 DR Pfam; PF00485; PRK.1.
 DR PRINTS; PR00478; PHRIBKINASE.
 DR PRINTS; PR00988; URIDINKINASE.
 KW Hypothetical protein; Transferrase; Kinase; ATP-binding.
 FT NP BIND 15 22 ATP (POTENTIAL).
 SQ SEQUENCE 229 AA; 26195 MW; 89CB21596A820FCC CRC64;

Query Match 35.3%; Score 512.5; DB 1; Length 229;
 Best Local Similarity 50.0%; Pred. No. 2e-35;
 Matches 105; Conservative 40; Mismatches 58; Indels 7; Gaps 4;

QY 23 PFLIGSGGTASGKSTVCEKIMELGQNEVEQ-RQRKVVILSDREFYKVLTAEOKAKALK 81
 DB 8 FLIIGVAGGTSCKSKSTIVERTIENINAMKQSRQDIIVHLSHSFYRELSEAEKILARE 67
 QY 82 GQYNFDPHPADNDMLHRTLKNIIVEGKTVFPPYDFTVTHSRLEPPTTVVPADVLFEGIL 141
 DB 68 GKFNFDPHPDQINPDILAETLQNMIDGKTVFIPKYMIT--SSNMGTVVEPAKVIIEGIL 126
 QY 142 VFSQEIQDMFHLRFVDTSDVRLSRR---LRDVRGRDLEQLIQYTFEVKPAPEEF 158
 DB 127 LLYDERVRKLSLTKLFVEKKAESRLNRLATYIRDRAP--LSITRYTEFKPAPEEF 184
 QY 199 CLPTKKYADYILPRGVNMAINLVQHIO 228
 DB 185 CRPTKKYADYILPRGADNHVATDLIAKNIQ 214

RESULT 9
 ID URK BACHD STANDARD; PRT; 211 AA.
 URK BACHD

AC 09KDB8;
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Uridine kinase (EC 2.7.1.48) (Uridine monophosphokinase) (Cytidine monophosphokinase).
 GN URK OR BH1275.
 OS Bacillus halodurans.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OC NCBI_TaxID=86665;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C-125 / JCM 9153;
 RX MEDLINE=20512582; PubMed=11058132;
 RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Maui N., Fujii F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S., Horikoshi K.;
 RL "Complete genome sequence of the alkaliphilic bacterium Bacillus RT halodurans and genomic sequence comparison with Bacillus subtilis.";
 CC -1- CATALYTIC ACTIVITY: ATP + uridine = ADP + UMP.
 CC -1- CATALYTIC ACTIVITY: ATP + cytidine = ADP + CMP.
 CC -1- PATHWAY: Pyrimidine salvage pathway.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AP001511; BAB04994.1; -.
 DR PIR; C83809; C83809.
 DR HAMAP; MF_00551; -; 1.
 DR InterPro; IPR003593; AAA_ATPase.
 DR InterPro; IPR006082; PRK.
 DR InterPro; IPR006083; PRK.
 DR InterPro; IPR00764; Uridine_kin.
 DR Pfam; PF00485; PRK.1.
 DR PRINTS; PR00478; PHRIBKINASE.
 DR PRINTS; PR00988; URIDINKINASE.
 DR TRANSFAM; TIGR00235; udk.1.
 KW Transferrase; Kinase; ATP-binding; Complete proteome.
 FT NP BIND 12 19 ATP (POTENTIAL).
 SQ SEQUENCE 211 AA; 24387 MW; C2AFA2CB0030520B CRC64;

	Query Match	32.6%	Score 473;	DB 1;	Length 211;
	Best Local Similarity	43.3%	Pred. No. 3	4e-32;	
	Matches	93;	Conservative	50;	Mismatches 56; Indels 16; Gaps 4
Oy	21 QRPPLIGVSGGTASGKSTVCSEKINELLQONEVEORQRKVILISODRFYK--VLTAEOKA	77			
Db	3 KRPIIIIGAGGTGGSGKTVAKEIF-----YQFNEKSIIVILEQDPAAYKDSQSLEERL	55			
Oy	78 KALGOVVFHPDPDAFNNDLMHRTLNIVEGKVEVPPTDFWHSRLPFTVVYPADNVILF	137			
Db	56 -----QNYIDPHLPADFNDLLIEHLHSLANGQAIEKPVIDYKLHTRSNVIYLPEKDVIIL	110			
Oy	138 EGIIVFYSGEIRDMENHRLPVDTSDVRLSRRVLDDR-RGRDLQLLYTTFFVKPAFE	196			
Db	111 EGILLIEDERLRREIMDIKLPFDVDADIRIIRRVADIERGERGTLESVIEQTKTVRPMHM	170			
Oy	197 EFCLPTKKCADVTIPRGVDNNVAIILVQHODIL 231				
Db	171 QFIETPKRYADVILIPEGQNRYVALDMLMTKTRAIL 205				

RESULT 10			
URK_CLOPE			
ID_URK_CLOPE	STANDARD:	PRT:	208 AA.
AC	Q8XJ16;		
DT	28-FEB-2003 (Rel. 41, Created)		
DT	28-FEB-2003 (Rel. 41, Last sequence update)		
DT	28-FEB-2003 (Rel. 41, Last annotation update)		
DE	Uridine kinase (EC 2.7.1.48) (Uridine monophosphokinase) (Cytidine monophosphokinase).		
DE	monophosphokinase).		
GN	UDK OR CPE1770.		
OS	Clostridium perfringens.		
OC	Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;		
OC	Clostridium.		
OX	NCBI_TaxID=1502;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=13 / Type A;		
RA	MEDLINE=21664373; PubMed=11792842;		
RA	Shimizu T., Ohtani K., Hirakawa H., Ohshima K., Yamashita A.,		
RA	Shiba T., Ogasawara N., Hattori M., Kuhara S., Hayashi H.;		
RT	"Complete genome sequence of Clostridium perfringens, an anaerobic		
RT	flesh-eater";		
RL	Proc. Natl. Acad. Sci. U.S.A. 99:996-1001(2002).		
CC	-1- CATALYTIC ACTIVITY: ATP + uridine = ADP + UMP.		
CC	-1- CATALYTIC ACTIVITY: ATP + cytidine = ADP + CMP.		
CC	-1- PATHWAY: Pyrimidine salvage pathway.		
CC	-1- SUBCELLULAR LOCATION: Cytoplasmic (by similarity).		
CC	-1- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/		
CC	or send an email to license@isb-sib.ch).		
CC	-----		
DR	EMBL; AP003191; BAB81476.1; -		
DR	HANAB; MF_00551; -; 1.		
DR	InterPro; IPR006082; PRK.		
DR	InterPro; IPR006083; PRK URK.		
DR	InterPro; IPR000764; Uridine_kin.		
DR	Pfam; PF00485; PRK: 1		
DR	PRINTS; PR00478; PHRBLKINASE.		
DR	PRINTS; PR0988; URIDINKINASE.		
DR	TIGRFAMs; TIGR00235; udk; 1.		
KW	transferase; Kinase; ATP-binding; Complete proteome.		
FT	NP BIND 11 18 ATP (POTENTIAL).		
SQ	SEQUENCE 208 AA; 23884 MW; 941BC368740DDE5E CMC64;		
Query Match	32.3%;	Score 468;	DB 1; Length 208;

[illegible]

ID	URL	LACIA	STANDARD:	PRT:	206 AA.
AC	09CE21;				
DT	28-FEB-2003	(Rel. 41, Created)			
DT	28-FEB-2003	(Rel. 41, Last sequence update)			
DT	28-FEB-2003	(Rel. 41, Last annotation update)			
DE	Uridine kinase (EC 2.7.1.48) (Uridine monophosphokinase) (Cytidine monophosphokinase).				
DE	Uridine kinase (EC 2.7.1.48) (Uridine monophosphokinase) (Cytidine monophosphokinase).				
GN	UDK OR LLI160.				
OS	Lactococcus lactis (subsp. lactis) (Streptococcus lactis).				
OC	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.				
OX	NCBI_TaxId=1360;				
RN	[1]				
RE	SEQUENCE FROM N.A.				
RC	STRAIN=LI1403;				
RX	MEDLINE=21235186; PubMed=11373471;				
RA	Boletín A., Winkler P., Manger S., Jallón O., Malarek K.,				
RA	Weissenbach J., Ehrlich S.D., Sorokin A.;				
RT	"The complete genome sequence of the lactic acid bacterium Lactococcus lactis ssp. lactis IL1403.";				
RL	Genome Res. 11:751-753(2001).				
CC	-1- CATALYTIC ACTIVITY: ATP + uridine = ADP + UMP.				
CC	-1- CATALYTIC ACTIVITY: ATP + cytidine = ADP + CMP.				
CC	-1- PATHWAY: Pyrimidine salvage pathway.				
CC	-1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).				
CC	-1- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb.ch/announce/isb-sib.ch).				
CC	or send an email to license@isb-sib.ch).				
DR	EMBL; AE006396; AAK05758.1; -				
DR	PIR; D86832; D86832.				
DR	HAMAP; MF_00551; -; 1.				
DR	InterPro; IPR006082; PRK.				
DR	InterPro; IPR006083; PRK.				
DR	InterPro; IPR000764; Uridine_kin.				
DR	Pfam; PF00485; PRK; 1.				
DR	PRINTS; PR00478; PHRINKINASE.				
DR	PRINTS; PR00988; URIDINKINASE.				
DR	TIGRFAMs; TIGR00235; udk; 1.				
KW	Transferase; Kinase; ATP-binding; Complete proteome.				
FT	NP BIND 11 18 ATP (POTENTIAL).				
SO	SEQUENCE 206 AA; 23740 MW; 8238F052037DBEE CRC64;				
Query Match	31.3%;	Score 454;	DB 1;	Length 206;	
Best Local Similarity	42.6%;	Pred. NO. 1.2e-30;			
Matches 92;	Conservative 43;	Mismatches 65;	Indels 16;	Gaps 5;	

```

QY 21 ORPFLIGVSGGTASGKSTVCCKIMELLGQNEVEQRKRVILSDRPFYK--VLTAROKA 77
D 2 KKTLLIGVTGASGKSTVSHALTF-SNE-----RIANIENHSYKQDSHLEFERT 54
QY 78 KALKQGVNFDHPDAFNDLMDHRTLKNIVEGKVEVPYDVTSHSLPPTTVVYPADVL 137
D 55 KT-----NYDHPAFNDLMDLLENKDKLNGKAVAVPPTDYASHRSRDTIDFKRQVIVIEGIF 109
QY 138 EGIIVFSQEI RDMFHLRFLPDTSDVRLSRVLRDV-RRGDLEQILITQYTTVKAFAFE 196
D 110 EGIIVLEDERLDRMDLMDIKIPDTDDVRIIRIRREDIERERGLDSDVITQYLDVAKPMYH 169
QY 197 EFCPLTKYADVIIPRGVDMNVAINLIQVHIQDIILN 232
D 170 QPIETKRYADVIIPRGVDMNVAINLIQVHIQDIILN 205

RESULT 12
URK_STRPM STANDARD; PRT; 207 AA.
AC 099TN8;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Uridine kinase (EC 2.7.1.48) (Uridine monophosphokinase) (Cytidine monophosphokinase)
GN UDK OR SA141611 OR SA1439 OR MW1561.
OS Staphylococcus aureus (strain Mu50 / ATCC 700699),
OS Staphylococcus aureus (strain N315), and
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=158878, 158879, 196620,
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=Mu50 / ATCC 700699, and N315;
RX MEDLINE=21311952, PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Iian J.-Q., Ito T.,
RA Kanamori M., Matsumaru H., Matuyama A., Murakami H., Hosoyana A.,
RA Mitutani U.I.Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hisekawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus aureus.";
RL Lancet 357:1225-1240 (2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MM2;
RX MEDLINE=22040717, PubMed=12044378;
RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
RA Nagai Y., Iwama N., Asano K., Natmi T., Kuroda H., Cui L.,
RA Yamamoto K., Hiramatsu K.;
RT "Genome and virulence determinants of high virulence community-acquired MRSA.";
RL Lancet 359:1819-1827 (2002).
CC -1- CATALYTIC ACTIVITY: ATP + uridine = ADP + UMP.
CC -1- CATALYTIC ACTIVITY: ATP + cytidine = ADP + CMP.
CC -1- PATHWAY: Pyrimidine salvage pathway.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC DR EMBL; AP003362; BAB57773.1; -
CC DR EMBL; AP003362; BAB57773.1; -
CC DR EMBL; AP003362; BAB57773.1; -

```

```

DR EMBL; AP004827; BAB95426.1; -.
DR PIR; B69943; B69943.
DR HAMAP; MF_00551; 1.
DR InterPro; IPR006082; PRK.
DR InterPro; IPR006083; PRK URK.
DR InterPro; IPR000764; Uridine_kin.
DR Pfam; PF00485; PRK; 1.
DR PRINTS; PR00478; PHRIELKINASE.
DR PRINTS; PR00988; URIDINKINASE.
DR TIGRFAMs; TIGR00235; udk; 1.
KW Transferase; Kinase; ATP-binding; Complete proteome.
FT NP BIND 11
FT MP BIND 18
SQ SEQUENCE 207 AA; 23505 MW; FD3A6D4E531DB8D CRC64;

Query Match 30.8%; Score 447; DB 1; Length 207;
Best Local Similarity 41.5%; Pred. No. 4.7e-30;
Matches 88; Conservative 50; Mismatches 58; Indels 16; Gaps 4;

QY 25 LIGVSGGTASGKSTVCCKIMELLGQNEVEQRKRVILSDRPFYK--KYLTAOKAKAK 81
D 6 IIGIAGSGSGKSTVINEIKNL-----EGSHVALLADQYVYKQKHLTPERLET-- 56
QY 82 GQYNFDPHDAFNDLMDHRTLKNIVEGKVEVPYDVTSHSLPPTTVVYPADVLFEGL 141
D 57 ---NYDHPAFNDLMDLLENKDKLNGKAVAVPPTDYASHRSRDTIDFKRQVIVIEGIF 113
QY 142 VFISQEI RDMFHLRFLPDTSDVRLSRVLRDV-RRGDLEQILITQYTTVKAFAFEFCL 200
D 114 ALBNKVLDRMDMDKVIYDADADRIIRLRDTRKGRSDSVINQYLSVRPMHQFIE 173
QY 201 PTKKYADVIIPRGVDMNVAINLIQVHIQDIILN 232
D 174 PTKKYADVIIPRGVDMNVAINLIQVHIQDIILN 205

RESULT 13
URK_STRPM STANDARD; PRT; 208 AA.
AC 089DPR;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Uridine kinase (EC 2.7.1.48) (Uridine monophosphokinase) (Cytidine monophosphokinase)
GN UDK OR SBYM18_1380.
OS Streptococcus pyogenes (serotype M18).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=186103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MGAS8232 / Serotype M18;
RX MEDLINE=21927593; PubMed=11917108;
RA Smoot J.C., Barbican K.D., Van Gompel J.J., Smoot L.M., Chaussee M.S.,
RA Sylva G.L., Sturdevant D.E., Ricklefs S.M., Porcella S.F.,
RA Parkins L.D., Beres S.B., Campbell D.S., Smith T.M., Zhang Q.,
RA Kapur V., Daly J.A., Veasy L.G., Musser J.M.;
RT "Genome sequence and comparative microarray analysis of serotype M18 group A Streptococcus strains associated with acute rheumatic fever outbreaks.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:4668-4673 (2002).
CC -1- CATALYTIC ACTIVITY: ATP + uridine = ADP + UMP.
CC -1- CATALYTIC ACTIVITY: ATP + cytidine = ADP + CMP.
CC -1- PATHWAY: Pyrimidine salvage pathway.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

```

```

CC or send an email to license@isb-sib.ch).
-----
CC EMBL: AE010058; AAL97975.1; -.
DR HAMAP; MF_00551; -.
DR InterPro; IPR006082; PRK.
DR InterPro; IPR006083; PRK.
DR InterPro; IPR00764; Uridine_kin.
DR Pfam; PF00485; PRK; 1.
DR PRINTS; PR00478; PHRIBKINASE.
DR PRINTS; PR00988; URIDINKINASE.
DR TIGRPFMS; TIGR00235; udk; 1.
DR Transferrase; Kinase; ATP-binding; Complete proteome.
FT NP BIND 12 19 ATP (POTENTIAL).
SQ SEQUENCE 208 AA; 23850 MW; BIA76ABF2C5AF511 CRC64;

Query Match 30.3%; Score 439.5; DB 1; Length 208;
Best Local Similarity 39.7%; Pred. No. 2e-29;
Matches 87; Conservative 51; Mismatches 64; Indels 17; Gaps 5;

QY 21 ORPRLIGVSGGTASGKSTVECKIMELLGQNEVEGRQKVVILSODRFYK---VLTAEQKA 77
DB 3 KKPILIGVTGSGGGKTSVSRALIDSF-----PNARIMTIOHDSYVKDQSHMSFEERV 55
QY 78 KALKGOYNFDPHDPFNDLHRTLNKIVEKTVVEPTVDFVTHSRLEPTTVVYPADVLF 137
DB 56 KT-----NYDHPPLAFPTDFMIQQLKELLAGRPVDIPYDKKHTRSMTTFRQDPQDVIIY 110
QY 138 EGLIVFYSQETIRDMFHLRFVDTSDVRLSRVLRD-VRRGRDLEQILTYTTFFVKPAFE 196
DB 111 EGLIVLEDELRDLMDIKLFVDTDDIRIRIRIKRDMERGRSLIESIIDQYTSVVKPMYH 170
QY 197 EFCLPTKKYADVIRPGVNMVAIINLVQHIDILNGDI 235
DB 171 QFIEPSKRYADVIIPBGVSNVAIDVINSKIASIL-GEV 208

RESULT 14
URK_STRPY STANDARD; PRT; 208 AA.
AC Q99Z70;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Uridine kinase (EC 2.7.1.48) (Uridine monophosphokinase) (Cytidine
DE monophosphokinase).
GN UDK OR SP1368 OR SPY3_1042 OR SPS0818.
OS Streptococcus pyogenes, and
OS Streptococcus pyogenes (serotype M3).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314, 198466;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MGAS315 / Serotype M3;
RX MEDLINE=21132684; PubMed=11296296;
RA Ferrari J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,
RA Primeaux C., Seate S., Suvorov A.N., Kenon S., Lai H.S., Lin S.P.,
RA Qian Y., Uta H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White J.,
RA Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;
RA "Complete genome sequence of an M1 strain of Streptococcus pyogenes.";
RA Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).
RL [2]
RN SEQUENCE FROM N.A.
RC STRAIN=MGAS315 / Serotype M3;
RX MEDLINE=21133808; PubMed=12122206;
RA Bares S.B., Sylva G.L., Barbhan K.D., Lei B., Hoff J.S.,
RA Mammarella N.D., Liu M.-Y., Smoot J.C., Porcella S.F., Parkins L.D.,
RA Campbell D.S., Smith T.M., McCormick J.K., Leung D.Y.M.,
RA Schlievert P.M., Musser J.M.;
RA "Genome sequence of a serotype M3 strain of group A Streptococcus:
RA phage-encoded toxins, the high-virulence phenotype, and clone
RA emergence.";
RA Proc. Natl. Acad. Sci. U.S.A. 99:10078-10083(2002).

```

```

RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=SSI-1 / Serotype M3;
RA Nakagawa I., Kurokawa K., Nakata M., Tomiyasu Y., Yamashita A.,
RA Hayazaki K., Okahashi N., Kawabata S., Yasunaga T., Hattori M.,
RA Hayaishi H., Hamada S.;
RT "The genome of invasive Streptococcus pyogenes; a comparative analysis
RT of S. pyogenes SSI-1, SP170 and MGAS832.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: ATP + uridine = ADP + UMP.
CC -1- CATALYTIC ACTIVITY: ATP + cytidine = ADP + CMP.
CC -1- PATHWAY: Pyrimidine salvage pathway.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (by similarity).
CC -1- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AE006574; AAK34195.1; -.
CC EMBL: AE014157; AAM79649.1; -.
CC EMBL: AP005143; BAC63913.1; -.
CC HAMAP; MF_00551; -.
CC InterPro; IPR006082; PRK.
CC InterPro; IPR006083; PRK.
CC InterPro; IPR00764; Uridine_kin.
CC Pfam; PF00485; PRK; 1.
CC PRINTS; PR00478; PHRIBKINASE.
CC PRINTS; PR00988; URIDINKINASE.
CC TIGRPFMS; TIGR00235; udk; 1.
CC Transferrase; Kinase; ATP-binding; Complete proteome.
FT NP BIND 12 19 ATP (POTENTIAL).
SQ SEQUENCE 208 AA; 23816 MW; 510007D58B379DB6 CRC64;

Query Match 30.3%; Score 439.5; DB 1; Length 208;
Best Local Similarity 39.7%; Pred. No. 2e-29;
Matches 87; Conservative 51; Mismatches 64; Indels 17; Gaps 5;

QY 21 ORPRLIGVSGGTASGKSTVECKIMELLGQNEVEGRQKVVILSODRFYK---VLTAEQKA 77
DB 3 KKPILIGVTGSGGGKTSVSRALIDSF-----PNARIMTIOHDSYVKDQSHMSFEERV 55
QY 78 KALKGOYNFDPHDPFNDLHRTLNKIVEKTVVEPTVDFVTHSRLEPTTVVYPADVLF 137
DB 56 KT-----NYDHPPLAFPTDFMIQQLKELLAGRPVDIPYDKKHTRSMTTFRQDPQDVIIY 110
QY 138 EGLIVFYSQETIRDMFHLRFVDTSDVRLSRVLRD-VRRGRDLEQILTYTTFFVKPAFE 196
DB 111 EGLIVLEDELRDLMDIKLFVDTDDIRIRIRIKRDMERGRSLIESIIDQYTSVVKPMYH 170
QY 197 EFCLPTKKYADVIRPGVNMVAIINLVQHIDILNGDI 235
DB 171 QFIEPSKRYADVIIPBGVSNVAIDVINSKIASIL-GEV 208

RESULT 15
URK_STRPN STANDARD; PRT; 212 AA.
AC Q97Q77;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Uridine kinase (EC 2.7.1.48) (Uridine monophosphokinase) (Cytidine
DE monophosphokinase).
GN UDK OR SP1208 OR SPRI090.
OS Streptococcus pneumoniae, and
OS Streptococcus pneumoniae (strain ATCC BAA-255 / R6).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.

```

```
OX NCBI_TaxID=1313, 171101;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC BAA-334 / TIGR4;
RX MEDLINE=21357209; PubMed=11463916;
RA Tetschlin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,
RA Peterson S., Heidelberg J., DeBoy R.T., Haft D.H., Dodson R.J.,
RA Durkin A.S., Gwinn M., Kolonay J.F., Nelson W.C., Peterson J.D.,
RA Umayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,
RA Holtzapple E., Khouli H., Wolf A.M., Uterback T.R., Hansen C.L.,
RA McDonald L.A., Feldblyum T.V., Angiuoli S., Dickinson T., Hickey E.K.,
RA Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,
RA Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;
RT "Complete genome sequence of a virulent isolate of Streptococcus
RT pneumoniae";
RL Science 293:498-506(2001).
[2]
RN SEQUENCE FROM N.A.
RP STRAIN=ATCC BAA-255 / R6;
RX MEDLINE=21429245; PubMed=11544234;
RA Hoskins J., Alborn W.E. Jr., Arnold J., Blaszczak L.C., Burgett S.,
RA Dehoff B.S., Bacterem S.T., Fritz L., Fu D.-U., Fuller W., Geringer C.,
RA Gilmour R., Glass J.S., Khoja H., Kraft A.R., Lagace R.E.,
RA Leblanc D.J., Lee L.N., Lefkowitz E.J., Lu J., Matsushima P.,
RA Mcahren S.M., McHenry M., Mcleaster K., Mundy C.W., Nicas T.I.,
RA Norris F.H., O'Gara M., Peery R.B., Robertson G.T., Rocky P.,
RA Sun P.-M., Winkler M.E., Yang Y., Young-Bellido M., Zhao G.,
RA Zook C.A., Baltz R.H., Uekunas S.R., Kosteck P.R. Jr., Skatrud P.L.,
RA Glass J.I.;
RT "Genome of the bacterium Streptococcus pneumoniae strain R6.";
RL J. Bacteriol. 183:5709-5717(2001).
CC -1- CATALYTIC ACTIVITY: ATP + uridine = ADP + UMP.
CC -1- CATALYTIC ACTIVITY: ATP + cytidine = ADP + CMP.
CC -1- PATHWAY: Pyrimidine salvage pathway.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (by similarity).
CC -1- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; AB007421; AAK75315.1; -
DR EMBL; AB008482; AAK9893.1; -
DR PIR; A99008; A99008.
DR PIR; B95140; B95140.
DR TIGR; SP1208; -
DR HAMAP; MF_00551; -; 1.
DR InterPro; IPR006082; PRK.
DR InterPro; IPR006083; PRK_URK.
DR InterPro; IPR00764; Uridine_kin.
DR Pfam; PF00485; PRK; 1.
DR PRINTS; PRO0478; PHRIPLKINASE.
DR PRINTS; PRO0988; URIDINKINASE.
DR TIGRPFAMs; TIGR00235; urk; 1.
KW Transferase; Kinase; ATP-binding; Complete proteome.
FT NP_BIND 12 19 ATP (POTENTIAL).
SQ SEQUENCE 212 AA; 24466 MW; 47B405D377C70A8A CRC64;

Query March 30.2%; Score 438; DB 1; Length 212;
Best Local Similarity 41.1%; Pred. No. 2.7e-29;
Matches 88; Conservative 42; Mismatches 68; Indels 16; Gaps 4;

QY 22 RPFLIGVSGTASGKSTVCEKIMELGNEVEQRQRKVVILSDRFYK---VLTAEQKAK 78
DQ 4 RPIITGVGTGGSGGKTSVSRALSS-----HPDEKISMEIHDSYKXQSHLTFEERVK 56
QY 79 ALKQGVNFDHPDADNDIMHRTLKNIVEGKTVENPTVDFTVHSRLPETTVVYPADVVLFE 138
DQ 57 T-----NYDHPFADPTDLMIEQIKELLAGRPVDIPYDYTEHTRSSKTYRQEPQDVFIIVE 111
```

```
QY 139 GILVFYSQELRDMFHLRLFVDTSDSVRLSRVLADV--RRGRDLQQLTOYTTFKPAFEE 197
DQ 112 GILVLEDKRLRLMDIKIFVDTDDVRLIRIKRDMERGRSLDSVINQTLGVAKPMYHQ 171
QY 198 FCLPTKKYADVILIPRGVDMVAINLIVQHIQDIL 231
DQ 172 FIESTKRYADIVIPEGSVNTVAIDLTLTKIAKIL 205
```

Search completed: November 25, 2003, 08:06:16
Job time : 14 secs


```

A:Molecule type: DNA
A:Residues: 1-555 <Mili>
A:Cross-references: EMBL:Z69635; NID:G1200023; PIDN:CAA93462.1; GSPDB:GN00022; CESP:F19B
A:Experimental source: clone F19B6
C:Genetics:
A:Gene: CESP:F19B6.1b
A:Map position: 4
A:Introns: 35/1; 83/3; 201/3; 267/3; 395/2; 511/3

Query Match
Best Local Similarity 52.9%; Pred. No. 8.8e-39;
Matches 110; Conservative 44; Mismatches 46; Indels 8; Gaps 2;

Qy 21 QRPFLIGSGGTASGKSTVCEKIMELLGQNEVEORQRKVILISODRFYKULTAEQKAKL 80
Db 102 KHPFVIGVCGSGASGKTTVAKEIVERLG-----IPWVTLISMDSFYKULTPEIRAAH 154

Qy 81 KGQYFDPHPDAPNDLMHRTLKNIVEGKTEVPPTYDFVTHSRLEPPTVVYPADVVLFEGL 140
Db 155 ESRINFDPGNAPFDLDLYEVKRLREGKSVDPVYDFDTHSRDPSKMGADVLIPEGI 214

Qy 141 LVFYSQEIIRDMFHLRLFVDTSDVRLSRVLDRV-RCRDLEQILTYTTFFVKPAFEFC 199
Db 215 LAFDERIKNIMDKMFVDTIDGDLRLARRIVRDVTDGRDIDIGMEQYFTEVKPAFDKYI 274

Qy 200 LPTKKYADVILIPRGVDNMVAINLIYOH 227
Db 275 APCMDSADLVIPRGENDVAIDMIVONY 302

RESULT 3
T18629
hypothetical protein B0001.4 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000
C:Accession: T18629
R:Sims, M.
submitted to the EMBL Data Library, February 1996
A:Reference number: Z18999
A:Accession: T18629
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-229 <Mili>
A:Cross-references: EMBL:Z69634; PIDN:CAA93453.1; GSPDB:GN00022; CESP:B0001.4
A:Experimental source: clone B0001
C:Genetics:
A:Gene: CESP:B0001.4
A:Map position: 4
A:Introns: 21/3; 41/3; 108/2
C:Superfamily: uridine kinase

Query Match
Best Local Similarity 35.3%; Score 512.5; DB 2; Length 229;
Matches 105; Conservative 40; Mismatches 58; Indels 7; Gaps 4;

Qy 23 PFLIGVSGGTASGKSTVCEKIMELLGQNEVEQ-RQRKVILISODRFYKULTAEQKAKL 81
Db 8 PLLIGVAGGISTCGSTIYERILLENLNANAKOSGRQIDIVHLSHFRELSAEKILARE 67

Qy 82 GQYNFDPHPDAPNDLMHRTLKNIVEGKTEVPPTYDFVTHSRLEPPTVVYPADVVLFEGL 141
Db 68 GKFPFDPHPDAPNDLMHRTLKNIVEGKTEVPPTYDFVTHSRLEPPTVVYPADVVLFEGL 126

Qy 142 VFYSQEIIRDMFHLRLFVDTSDVRLSRV---LADVRGRDLEQILTYTTFFVKPAFEFC 198
Db 127 LLYDERVAKLSTKLTFEKNAESRLRNRLATYIRIDYRAP--LSIIROYTEFVAPAEFC 184

Qy 199 CLPTKKYADVILIPRGVDNMVAINLIYOH 228
Db 185 CRPTKKYADVILIPRGVDNMVAIDLIANKQL 214

RESULT 4

```

```

C83809
uridine kinase udk [imported] - Bacillus halodurans (strain C-125)
C:Species: Bacillus halodurans
C>Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C:Accession: C83809
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hirai
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and c
A:Reference number: A83650; MIMD:20512582; PMID:11058132
A:Reference number: C83809
A:Accession: C83809
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-211 <STO>
A:Cross-references: GB:AP001511; GB:BA000004; NID:G10173727; PIDN:BAB04994.1; GSPDB:GN001
A:Experimental source: strain C-125
C:Genetics:
A:Gene: udk
C:Superfamily: uridine kinase

Query Match
Best Local Similarity 32.6%; Score 473; DB 2; Length 211;
Matches 93; Conservative 50; Mismatches 56; Indels 16; Gaps 4;

Qy 21 QRPFLIGSGGTASGKSTVCEKIMELLGQNEVEORQRKVILISODRFYK---VLTAEQKA 77
Db 3 KRPIIIGVAGSGTSGKTTVAKEIF-----YQFNKSVLVLBOADYKDGQSLSEERL 55

Qy 78 KALKGQYFDPHPDAPNDLMHRTLKNIVEGKTEVPPTYDFVTHSRLEPPTVVYPADVVLFE 137
Db 56 ----QTNYPDHPAPFNDLDLIEHLHSLNGOAIIEKVVYDYKLTSTSENEVILVEPKDVIIL 110

Qy 138 EGLVYFSQEIIRDMFHLRLFVDTSDVRLSRVLDRV-RCRDLEQILTYTTFFVKPAFE 196
Db 111 EGLVLEDERLKLMDKLFVDTDADIRILIRRVDIRRGRKTLSEVQYTKVAPMEM 170

Qy 197 EFCLPTKKYADVILIPRGVDNMVAINLIYOHIDIL 231
Db 171 QFIEPTKKYADVILIPRGQNRVAIDMIVTKIRALI 205

RESULT 5
T41020
probable Uridine kinase - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C:Accession: T41020
R:Murphy, L.; Harris, D.; Wood, V.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, June 1998
A:Reference number: Z21965
A:Accession: T41020
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-454 <MUR>
A:Cross-references: EMBL:AL023860; PIDN:CAA19591.1; GSPDB:GN00068; SPDB:SPCC162.11c
A:Experimental source: strain 972h-; cosmid c162
C:Genetics:
A:Gene: SPDB:SPCC162.11c
A:Map position: 3

Query Match
Best Local Similarity 31.7%; Score 459; DB 2; Length 454;
Matches 92; Conservative 47; Mismatches 76; Indels 8; Gaps 2;

Qy 11 SPAPEADPHQRPFLIGVSGGTASGKSTVCEKIMELLGQNEVEORQRKVILISODRFYK 70
Db 9 SSNPTEPEPRKRVFRFGIAGPSGSGTSAQOLIVKALNPH-----VVIISLDSFYKS 61

Qy 71 LTAEQKAKALKGQYFDPHPDAPNDLMHRTLKNIVEGKTEVPPTYDFVTHSRLEPPTVVY 130
Db 62 LNAEQKRAFPNNDYDPSPAIDWDLFYKLELKQGRVVDIPIYSFNEHNRILPEPTNITLF 121

Qy 131 PADVLFEGILVYFSQEIIRDMFHLRLFVDTSDVRLSRVLDRV-RCRDLEQILTYTT 189

```


Db 122 GASIIIEGIFALYDEKIRSLDVSVELDTPDSVCLSRRLNRDINVRGRDVGVLQYKN 181
QY 190 FVNPAPFEECLPTKKYADVILPRGVDMVAIINLVQHIDILN 232
Db 182 FVNPSENFVRROLSTYDILVPRGRDNKALIDWVINFIRRTLS 224

RESULT 6

Db6832
uridine kinase (EC 2.7.1.48) [imported] - Lactococcus lactis subsp. lactis (strain IL140
C/Species: Lactococcus lactis subsp. lactis
C/Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001
C/Accession: D86832
R:Bohlochin, A.; Winkler, P.; Manger, S.; Jallion, O.; Malame, K.; Weissenbach, J.; Ehrlich
Genome Res. 11, 731-753, 2001
A/Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis s
A/Reference number: A86625; MUID:21235186; PMID:11337471
A/Accession: D86832
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-206 <STO>
A/Cross-references: GB:AE005176; PID:g12724672; PIDN:AAK05758.1; GSPDB:GN00146
A/Experimental source: strain IL1403
C/Genetics:
A/Gene: udk
C/Superfamily: uridine kinase
C/Keywords: phosphotransferase

Query Match 31.3%; Score 454; DB 2; Length 206;
Best Local Similarity 42.6%; Pred. No. 1.4e-30;
Matches 92; Conservative 43; Mismatches 65; Indels 16; Gaps 5;

QY 21 QRPPLIVSGGTASGKSTVEKIMELLGONEVEQRQKRVILSDRYK--VLTAEQKA 77
Db 2 KKLIIIGVTGSGSGKTSVSHAIETP-SNE-----RIMIBHDSYKQSHLTFEERT 54
QY 78 KATKGQVNFDPDADNDLNRITLKNIVEGKTVETPTVDVTSRLPETTVVYPADVLF 137
Db 55 KT-----NYDHPAFDIDYIAQLKELOYGRANDIPYDPAKTRSGEITRQEPVDVLI 109
QY 138 EGIIVFYSQIRMFHLRLFVDTSDVRLSRVLRDY-RGRDLEQILTYTTFFVKPAFE 196
Db 110 EGIIVFEDERLDMDKIFVDDDDVRIIRIRIDIEEGRTLDSTVITQYLVAKVMYH 169
QY 197 EFCPLTKKYADVILPRGVDMVAIINLVQHIDILN 232
Db 170 QFIEPTKRYADVIIPRGSNTVGVDDITTKIASILN 205

RESULT 7

Db89943
uridine kinase [imported] - Staphylococcus aureus (strain N315)
C/Species: Staphylococcus aureus
C/Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C/Accession: B89943
R:Kuroda, M.; Ohka, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cai, L.; Oguc
ma, A.; Mizutani-Oi, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
C.; Shibata, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A/Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A/Reference number: A89758; MUID:21311952; PMID:11418146
A/Accession: B89943
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-207 <KUR>
A/Cross-references: GB:BA000018; PID:g13701409; PIDN:BA842703.1; GSPDB:GN00149
A/Experimental source: strain N315
C/Genetics:
A/Gene: udk
C/Superfamily: uridine kinase

Query Match 30.8%; Score 447; DB 2; Length 207;
Best Local Similarity 41.5%; Pred. No. 5.4e-30;

Matches 88; Conservative 50; Mismatches 58; Indels 16; Gaps 4;

QY 25 LIGVSGGTASGKSTVEKIMELLGONEVEQRQKRVILSDRYK--VLTAEQKAK 81
Db 6 IIGIAGSGSGKTVTNEIMKNL-----EGSHVALLAQDYKKQKHLTFDERLET-- 56
QY 82 GQYNFDPDADNDLNRITLKNIVEGKTVETPTVDVTSRLPETTVVYPADVLFEGIL 141
Db 57 ---NYDHPAFDNDLIEMLKDLKNGKAVEPPIYDASHITRSDITIDPKKVIYIEGIF 113
QY 142 VFYSQIRMFHLRLFVDTSDVRLSRVLRDY-RGRDLEQILTYTTFFVKPAFEFCL 200
Db 114 ALENKVLRLDMMDVKIYVDIDADRLIRLRLTRDTKERGRMSDVINGYLSVRPMHQFIE 173
QY 201 PTKKYADVILPRGVDMVAIINLVQHIDILN 232
Db 174 PTKKYADVIIPRGSNTVGVDDITTKIQSLVS 205

RESULT 8

Db95140
uridine kinase [imported] - Streptococcus pneumoniae (strain TIGR4)
C/Species: Streptococcus pneumoniae
C/Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 24-Aug-2001
C/Accession: B95140
R:Tetzelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid
on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple, I
nson, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A/Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
A/Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
A/Reference number: A95000; MUID:21357209; PMID:11463916
A/Accession: B95140
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-212 <KUR>
A/Cross-references: GB:AE005672; PIDN:AAK75315.1; PID:g14972689; GSPDB:GN00164; TIGR:SP46
A/Experimental source: strain TIGR4
C/Genetics:
A/Gene: SP1208
C/Superfamily: uridine kinase

Query Match 30.2%; Score 438; DB 2; Length 212;
Best Local Similarity 41.1%; Pred. No. 3.1e-29;
Matches 88; Conservative 42; Mismatches 68; Indels 16; Gaps 4;

QY 22 RPLILVSGGTASGKSTVEKIMELLGONEVEQRQKRVILSDRYK--VLTAEQKAK 78
Db 4 RPIIIGVTGSGSGKTSVSRALIS-----HPPDEKIMIHDSYKQSHLTFEERYK 56
QY 79 ALKGQVNFDPDADNDLNRITLKNIVEGKTVETPTVDVTSRLPETTVVYPADVLF 138
Db 57 T-----NYDHPAFDIDLIEQIKELAGRPVIDPYDTEHTRSSKTRQEPDVFIVE 111
QY 139 GIVFYSQIRMFHLRLFVDTSDVRLSRVLRDY-RGRDLEQILTYTTFFVKPAFE 197
Db 112 GIVFIEDKRLDMDKIFVDDDDVRIIRIRIDIEEGRTLDSTVITQYLVAKVMYH 171
QY 198 FCLPTKKYADVILPRGVDMVAIINLVQHIDILN 231
Db 172 FIESTKRYADVIIPRGSNTVGVDDITTKIAKIL 205

RESULT 9

DbA99008
uridine kinase (EC 2.7.1.48) [imported] - Streptococcus pneumoniae (strain Re)
C/Species: Streptococcus pneumoniae
C/Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 02-Nov-2001
C/Accession: A99008
R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; E
e, R.; Leblanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M
Y, P.; Sun, P.M.; Winkler, M.E.
J. Bacteriol. 183, 5709-5717, 2001


```
Db 3 KNPVVIAGSGSGKTSVTRSIYE-----QFKGSHILMIQODLVYKQSHLPFEERL 55
Qy 78 KALGQGNFHPDPAFNDMLHRTLKNIVEGKTVVPTPDFTHSRLPPTTVYPADVYLF 137
Db 56 NT-----NYDHPAFDNDYLLIEHIQDILNYPRIKPIYDYKLTARSEETVYVEPKDYIL 110
Qy 138 EGIIVFYSQIEIRDMFHLRFVDTSDVRLSRVLRDV-RRGRDLEQILTYTTVFYKPAFE 196
Db 111 EGIIVLEDRDLMDIRLFDVTDADRLIRIRKREDIRKERSIDSVTEQVSVVRPMAN 170
Qy 197 EFCLEPTKRYADVILIPRGVDNMVAIINLIVQHIQDIL 231
Db 171 QFVEPTKRYADIIIPGGQGNHVAIDLMTVKIQITIL 205

RESULT 12
F84277
uridine kinase [imported] - Halobacterium sp. NRC-1
C/Species: Halobacterium sp. NRC-1
C/Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 17-May-2002
C/Accession: F84277
R/Ref: W.V.; Kennedy, S.P.; Mahairas, G.G.; Bergquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.;
Leithausen, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablo
Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A/Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li
A/Title: Genome sequence of Halobacterium species NRC-1.
A/Reference number: A84160; MUID:20504483; PMID:11016950
A/Accession: F84277
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-248 <STO>
A/Cross-references: GB:AE004437; NID:g10580750; PIDN:AA619586.1; GSPDB:GN00138
C/Genetics:
A/Gene: urk
C/Superfamily: uridine kinase

Query Match 28.8%; Score 417; DB 2; Length 248;
Best Local Similarity 39.7%; Pred. No. 2.2e-27;
Matches 83; Conservative 50; Mismatches 58; Indels 18; Gaps 5;

Qy 24 FLIGVSGTASGKSTVCEKIMELLGONEVEQORQRKVVILSQDRFYKVL--AEQAKAL 80
Db 29 FAIGTAGTGAGKTTVAKEITDNGES-----ATLIPLDNYEDLSRPFERRANA- 79
Qy 81 KQGNFHPDPAFNDMLHRTLKNIVEGKTVVPTPDFTHSRLPPTTVYPADVYLFEGI 140
Db 80 ----NYDHPAFEFWELRLTHIDALLSGQSIEMPGYDEFERHVRKADRVVVEPTDVIYLEGI 135
Qy 141 LVFYSQIEIRDMFHLRFVDTSDVRLSRVLRDV-RRGRDLEQILTYTTVFYKPAFEFC 199
Db 136 LAISDETVNDMLDHIYETDADVRLIRIRKREDIRKERSIDSVTEQVSVVRPMAN 195
Qy 200 LPTKRYADVILIPRGVDNMVAIINLIVQHIQ 228
Db 196 EPTKRYADIIIPGGQGNHVAIDLMTVKIQ 223

RESULT 13
D86582
uridine kinase [imported] - Chlamydia pneumoniae (strain J138)
C/Species: Chlamydia pneumoniae, Chlamydia pneumoniae
C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001
C/Accession: D86582
R/Ref: R.; Shirai, M.; Hiraoka, H.; Kinoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; I
Nucleic Acids Res. 28, 2311-2314, 2000
A/Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.
A/Reference number: A86491; MUID:20330349; PMID:10871362
A/Accession: D86582
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-222 <STO>
A/Cross-references: GB:BA000008; NID:g8979107; PIDN:BA98942.1; GSPDB:GN00142
```

```
A/Experimental source: strain J138
C/Genetics:
A/Gene: CPJ0735
C/Superfamily: uridine kinase

Query Match 28.4%; Score 412; DB 2; Length 222;
Best Local Similarity 38.4%; Pred. No. 4.9e-27;
Matches 84; Conservative 54; Mismatches 59; Indels 22; Gaps 6;

Qy 25 LIQVSGTASGKSTVCEKIMELLGONEVEQORQRKVVILSQDRFYK--VLTAAQAKALK 81
Db 8 LIITGSGGAGKTTLQNIKEIFGED-----VAVICQDNYKYKRSHTPEERANLI- 58
Qy 82 GQYNFHPDPAFNDMLHRTLKNIVEGKTVVPTPDFTHSRLPPTTVYPADVYLFEGI 140
Db 59 ----WDHPDPAFNDLLISDIKRLKNEIYQAPVDFVLGNRSKTELETTPSKVILVEGI 114
Qy 141 LVFYSQIEIRDMFHLRFVDTSDVRLSRVLRDV-RRGRDLEQILTYTTVFYKPAFEFC 199
Db 115 LVFENQELRDLMDIRLFDVTDADRLIRIRKREDIRKERSIDSVTEQVSVVRPMAN 174
Qy 200 LPTKRYADVILIPRGVDNMVAIINLIVQ----HIQDILNGD 234
Db 175 EPTKRYADIIIPGGQGNHVAIDLMTVKIQITIL 205

RESULT 14
E72041
uridine kinase CP0011 [imported] - Chlamydia pneumoniae (strains CWL029 and AR39)
C/Species: Chlamydia pneumoniae, Chlamydia pneumoniae
C/Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-May-2000
C/Accession: E72041; E81622
R/Ref: R.; Kaiman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.;
Nature Genet. 21, 385-389, 1999
A/Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.
A/Reference number: A72000; MUID:99206606; PMID:10192388
A/Accession: E72041
A/Molecule type: DNA
A/Residues: 1-222 <ARN>
A/Cross-references: GB:AE001655; GB:AE001363; NID:g4377039; PIDN:AA018874.1; PID:g4377043
A/Experimental source: strain CWL029
R/Ref: R.; Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey, J.;
C.; Dodson, R.; Gwinn, M.; Nelson, W.; Deboy, R.; Kolonay, J.; McClarty, G.; Salzberg,
Nucleic Acids Res. 28, 1397-1406, 2000
A/Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
A/Reference number: A81500; MUID:20150255; PMID:10684935
A/Accession: E81622
A/Molecule type: DNA
A/Residues: 1-222 <REA>
A/Cross-references: GB:AE002165; GB:AE002161; NID:g7188948; PIDN:AA637907.1; PID:g7188950
A/Experimental source: strain AR39, HL cells
C/Genetics:
A/Gene: CP00735; CP0011
C/Superfamily: uridine kinase

Query Match 28.4%; Score 412; DB 2; Length 222;
Best Local Similarity 38.4%; Pred. No. 4.9e-27;
Matches 84; Conservative 54; Mismatches 59; Indels 22; Gaps 6;

Qy 25 LIQVSGTASGKSTVCEKIMELLGONEVEQORQRKVVILSQDRFYK--VLTAAQAKALK 81
Db 8 LIITGSGGAGKTTLQNIKEIFGED-----VAVICQDNYKYKRSHTPEERANLI- 58
Qy 82 GQYNFHPDPAFNDMLHRTLKNIVEGKTVVPTPDFTHSRLPPTTVYPADVYLFEGI 140
Db 59 ----WDHPDPAFNDLLISDIKRLKNEIYQAPVDFVLGNRSKTELETTPSKVILVEGI 114
Qy 141 LVFYSQIEIRDMFHLRFVDTSDVRLSRVLRDV-RRGRDLEQILTYTTVFYKPAFEFC 199
Db 115 LVFENQELRDLMDIRLFDVTDADRLIRIRKREDIRKERSIDSVTEQVSVVRPMAN 174
Qy 200 LPTKRYADVILIPRGVDNMVAIINLIVQ----HIQDILNGD 234
Db 175 EPTKRYADIIIPGGQGNHVAIDLMTVKIQITIL 205
```

Db 175 EPTKRYADIIVHGNVQNVNTIISQIKIKNLENALES2D 213

RESULT 15

G70101
 uridine kinase (udk) homolog - Lyme disease spirochete
 C:Species: Borrelia burgdorferi (Lyme disease spirochete)
 C:Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 17-Nov-2000
 C:Accession: G70101; T46979
 R:Fraser, C.M.; Caetiens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White
 son, D.; Peterson, J.; Kervage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt,
 ; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
 Nature 390, 580-586, 1997
 A:Authors: Smith, H.O.; Venter, J.C.
 A:Title: Genomic sequence of a Lyme disease spirochete. Borrelia burgdorferi.
 A:Reference number: A70100; PMID:98065943; PMID:9403685
 A:Accession: G70101
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-207 <KLE>
 A:Cross-references: GB:AE001115; GB:AE000783; NID:g2687879; PIDN:AA066392.1; PID:g268788
 A:Experimental source: strain B31
 R:Boursaux-Bude, C.; Margalla, D.; Gilles, A.M.; Barzu, O.; Girons, I.S.
 FEMS Microbiol. Lett. 151, 257-261, 1997
 A:Title: Borrelia burgdorferi uridine kinase: an enzyme of the pyrimidine salvage pathwa
 A:Reference number: 224339; MUID:97372541; PMID:9228761
 A:Accession: T46979
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 2-207 <BOU>
 A:Cross-references: EMBL:X97449; NID:g1405440; PIDN:CAA6081.1; PID:g1405442
 A:Experimental source: strain HB19, sep. senso stricto
 C:Genetics:
 A:Gene: udk
 C:Superfamily: uridine kinase

Query Match 28.3%; Score 410.5; DB 2; Length 207;
 Best Local Similarity 36.9%; Pred. No. 6e-27;
 Matches 79; Conservative 49; Mismatches 59; Indels 27; Gaps 4;
 QY 25 LIGVSGTASGKSTYCEKIMELGONEVEQORQKVIIISDRFYKVLTAEQKAKALKGQY 84
 Db 5 IIGISGSGSGKSTVVSKEISFIEP-----FVLISQDNYYKSV-----GDY 45
 QY 85 -----NEHHPDAFNDNLMRTLKNIVEGKTVSEVPTVDPFTHSRILPETTVVPADVVL 137
 Db 46 EHERSKVNFDPDAFDNNLFTFHLKNLKKNSPIDMPLDFTNKKQKLVLPVPTVIV 105
 QY 138 EGIIVFYSQELRDMFHLRLFVDTSDVRLSRVLADV-RRGRDLQQLITQYTFVKPAFE 196
 Db 106 EGINIFVERVRNLDLKIYIDTPNDIRFIRLRDLISKGRTVESVIDQYLTNTRMGY 165
 QY 197 EFCUPTKRYADVIIPRGVDNVAIIVQHIQDI 230
 Db 166 RFIPTKRYADLIIPGGHNDKALYVLSTFLKSL 199

Search completed: November 25, 2003, 08:07:26
 Job time : 22 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2003, 08:05:10 ; Search time 31 Seconds
(without alignments)
1648.090 Million cell updates/sec

Title: US-09-896-522-2
Perfect score: 1450
Sequence: 1 MASAGEDESPAPADRP...HPGMLTSGKSHLESSRPH 277

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 673684 seqs, 18443283 residues

Total number of hits satisfying chosen parameters: 673684

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Published Applications AA:*
1: /cgn2_6/ptodata/1/pubppa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubppa/US06_PUBCOMB.pep:*
3: /cgn2_6/ptodata/1/pubppa/US05_PUBCOMB.pep:*
4: /cgn2_6/ptodata/1/pubppa/US04_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubppa/US03_PUBCOMB.pep:*
6: /cgn2_6/ptodata/1/pubppa/US02_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubppa/US01_PUBCOMB.pep:*
8: /cgn2_6/ptodata/1/pubppa/US00_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubppa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubppa/US08_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubppa/US07_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubppa/US06_PUBCOMB.pep:*
13: /cgn2_6/ptodata/1/pubppa/US05_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubppa/US04_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubppa/US03_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubppa/US02_PUBCOMB.pep:*
17: /cgn2_6/ptodata/1/pubppa/US01_PUBCOMB.pep:*
18: /cgn2_6/ptodata/1/pubppa/US00_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1450	100.0	277	US-09-896-522-2	Sequence 2, Appl1
2	896	61.8	337	US-09-925-300-1160	Sequence 1160, Ap
3	428.5	29.6	125	US-09-896-522-5	Sequence 5, Appl1
4	327	22.6	60	US-10-029-386-28765	Sequence 28765, A
5	320.5	22.1	124	US-09-896-522-6	Sequence 6, Appl1
6	276	19.0	231	US-09-896-522-4	Sequence 4, Appl1
7	154.5	10.7	312	US-09-738-626-4609	Sequence 4609, Ap
8	149	10.3	339	US-10-156-761-12495	Sequence 12495, A
9	119.5	8.2	143	US-09-734-569-88	Sequence 88, Appl1
10	97	6.7	216	US-10-301-822-217	Sequence 217, App
11	95	6.6	251	US-10-301-822-220	Sequence 220, App
12	95	6.6	257	US-10-106-698-6128	Sequence 6128, Ap
13	95	6.6	267	US-10-301-822-215	Sequence 215, App
14	93	6.4	210	US-09-769-744A-64	Sequence 64, Appl1
15	93	6.4	433	US-09-815-242-5691	Sequence 5691, Ap

16	93	6.4	433	9	US-09-815-242-12690	Sequence 12690, A
17	91	6.3	234	15	US-10-156-761-10887	Sequence 10887, A
18	88	6.1	674	9	US-09-765-272-200	Sequence 200, App
19	87	6.0	457	12	US-09-896-522-394	Sequence 394, App
20	87	6.0	865	9	US-09-815-242-11364	Sequence 11364, A
21	87	6.0	910	12	US-09-896-186B-16	Sequence 16, Appl1
22	86	5.9	457	9	US-09-815-242-11616	Sequence 11616, A
23	84.5	5.8	252	9	US-09-925-301-1273	Sequence 1273, Ap
24	84.5	5.8	508	9	US-09-815-242-11555	Sequence 11555, A
25	84	5.8	600	11	US-09-893-519A-71	Sequence 71, Appl1
26	82.5	5.7	214	12	US-10-032-585-7414	Sequence 7414, Ap
27	82.5	5.7	231	12	US-10-220-381-17	Sequence 17, Appl1
28	82	5.7	665	15	US-10-216-556-18	Sequence 18, Appl1
29	81.5	5.6	508	12	US-09-883-227-570	Sequence 570, App
30	81	5.6	213	15	US-10-156-761-13762	Sequence 13762, A
31	81	5.6	245	15	US-10-156-761-11353	Sequence 11353, A
32	80.5	5.6	722	8	US-08-817-832B-32	Sequence 32, Appl1
33	80	5.5	645	10	US-09-919-172-41	Sequence 41, Appl1
34	80	5.5	744	10	US-09-835-081-4	Sequence 4, Appl1
35	80	5.5	877	9	US-09-815-242-4876	Sequence 4876, Ap
36	80	5.5	880	9	US-09-815-242-10675	Sequence 10675, A
37	79.5	5.5	231	9	US-09-815-242-5760	Sequence 5760, App
38	79.5	5.5	744	9	US-09-815-242-5473	Sequence 5473, Ap
39	79.5	5.5	744	9	US-09-815-242-12346	Sequence 12346, A
40	79	5.4	1070	12	US-10-032-585-7389	Sequence 7389, Ap
41	78.5	5.4	414	15	US-10-119-651-6	Sequence 6, Appl1
42	78.5	5.4	610	9	US-09-815-242-10414	Sequence 10414, A
43	78.5	5.4	691	10	US-09-919-585-6	Sequence 6, Appl1
44	78.5	5.4	724	10	US-09-919-585-9	Sequence 9, Appl1
45	78.5	5.4	730	15	US-10-217-371-6	Sequence 6, Appl1

ALIGNMENTS

RESULT 1
US-09-896-522-2
; Sequence 2, Application US/09896522
; Patent No. US20020055161A1
; GENERAL INFORMATION:
; APPLICANT: Glucksmann, Maria A.
; TITLE OF INVENTION: 57658, A NOVEL HUMAN URIDINE KINASE AND
; TITLE OF INVENTION: US85 THEREOF
; FILE REFERENCE: 381552001700
; CURRENT APPLICATION NUMBER: US/09/896,522
; CURRENT FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: 60/216,503
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 277
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-896-522-2

Query Match 100.0%; Score 1450; DB 9; Length 277;
Best Local Similarity 100.0%; Pred. No. 5.7e-151;
Matches 277; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MASAGEDESPAPADRP...PFLIGVSGTASGTCCKIMELGONEVQRKRV 60
DB 1 MASAGEDESPAPADRP...PFLIGVSGTASGTCCKIMELGONEVQRKRV 60
QY 1 MASAGEDESPAPADRP...PFLIGVSGTASGTCCKIMELGONEVQRKRV 120
DB 1 MASAGEDESPAPADRP...PFLIGVSGTASGTCCKIMELGONEVQRKRV 120
QY 1 MASAGEDESPAPADRP...PFLIGVSGTASGTCCKIMELGONEVQRKRV 180
DB 1 MASAGEDESPAPADRP...PFLIGVSGTASGTCCKIMELGONEVQRKRV 180
QY 1 MASAGEDESPAPADRP...PFLIGVSGTASGTCCKIMELGONEVQRKRV 240
DB 1 MASAGEDESPAPADRP...PFLIGVSGTASGTCCKIMELGONEVQRKRV 240

Db 181 EQLITOTTTTPKPAFEFECFLPTKCYADVIIPRGDNNVALNIQIIDLINGDICKMR 240
QY 241 GGSNGRSYKRTFSPGDHPGMLTSGKSHLESSSRPH 277
Db 241 GGSNGRSYKRTFSPGDHPGMLTSGKSHLESSSRPH 277

RESULT 2
US-09-925-300-1160
; Sequence 1160, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben,
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1160
; LENGTH: 337
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (38)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (46)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (155)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (169)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-300-1160

Query Match 61.8%; Score 896; DB 10; Length 337;
Best Local Similarity 62.3%; Pred. No. 7,6e-90;

Matches 188; Conservative 28; Mismatches 40; Indels 46; Gaps 6;

QY 2 ASAGGE-----DCSPAPRADR--PHQPFLLGVSGGTASGK 36
Db 56 AGAGGERASVRTSGRGKGNHNGORADPAEPAAQRRRLPYRR-----HGTTASGK 109
QY 37 STVCEKIMELLQNEVEQORQKVILSQDRFYKVLTAQKAKALKGQYNPDHPDAPNDL 96
Db 110 SSVCAKIVQLLQNEVEVQKQVILSQDSFYRLTSEQKAKALKQFNPDHPDAPDNEX 169
QY 97 MHTLKNIIVEKTVETPYDVTSLRSLBETTVVPAVVLFEGILVYSQEIEMFHLRL 156
Db 170 ILKTLKEITBEKTVQIIPYDVSLSRKEETVTVPAVVLFEGILVYSQEIEMFHLRL 229
QY 157 FVDTDSVRLSRVLRADV-RRGRLDQIILTYTTFVKPAFEFCLPTKKYADVIIPPGVD 215
Db 230 FVDTDSVRLSRVLRADV-RRGRLDQIILTYTTFVKPAFEFCLPTKKYADVIIPPGVD 289
QY 216 NMVAINIIVQHIQIIDLINGDICKMRGSGNGRSYKRTFSPGDHPGMLTSGKSHLESSSR 275
Db 290 NMVAINIIVQHIQIIDLINGDICKMRGSGNGRSYKRTFSPGDHPGMLTSGKSHLESSSR 335
QY 276 PH 277
Db 336 PH 337

RESULT 3
US-09-896-522-5
; Sequence 5, Application US/09896522
; Patent No. US20020055161A1
; GENERAL INFORMATION:
; APPLICANT: Gluckemann, Maria A.
; TITLE OF INVENTION: 57658, A NOVEL HUMAN URIDINE KINASE AND
; TITLE OF INVENTION: US8, THEREOF
; FILE REFERENCE: 381552001700
; CURRENT APPLICATION NUMBER: US/09/896,522
; CURRENT FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: 60/216,503
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 125
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Consensus amino acid sequence
US-09-896-522-5

Query Match 29.6%; Score 428.5; DB 9; Length 125;
Best Local Similarity 67.2%; Pred. No. 5e-39;
Matches 84; Conservative 12; Mismatches 28; Indels 1; Gaps 1;

QY 154 LRIFVDTSDVRLSRVLRADV-RRGRLDQIILTYTTFVKPAFEFCLPTKKYADVIIPR 212
Db 1 LKIFVDTSDVRLSRVLRADV-RRGRLDQIILTYTTFVKPAFEFCLPTKKYADVIIPR 60
QY 213 GVNMMVAINIIVQHIQIIDLINGDICKMRGSGNGRSYKRTFSPGDHPGMLTSGKSHLESS 272
Db 61 GSDNHVAIDLIVQHIQIIDLINGDICKMRGSGNGRSYKRTFSPGDHPGMLTSGKSHLESS 120
QY 273 SSRPH 277
Db 121 SSRPH 125

RESULT 4

US-10-029-386-28765
; Sequence 28765, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Hanzel, David R.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR GI
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: ABOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 28765
; LENGTH: 60
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR9.1
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.5
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.8
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.6
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.7
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.5
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.7
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.3
; OTHER INFORMATION: SWISSPROT HIT: P52623, EVALU8 8.00e-26
US-10-029-386-28765

Query Match 22.6%; Score 327; DB 12; Length 60;

RESULT 8

```
US-10-156-761-12495
; Sequence 12495, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156, 761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 12495
; LENGTH: 329
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-12495
```

```
Query Match 10.3%; Score 149; DB 15; Length 329;
Best Local Similarity 27.1%; Pred. No. 1.3e-07;
Matches 62; Conservative 36; Mismatches 89; Indels 40; Gaps 11;
```

```
QY 23 PELGVSGGTASGKSTVCEKIMELGQNEVEQRQRKVVLSQDFYKVLTAEQAKLKG 82
DB 100 PFGVAGSVAVGKSTVAFRLQLLSRWPEHPH---VELVTTDGL-LPTKELEARGLMS 155
QY 83 QYNVDHDAFNDMDHRTLNKIVEGK-TVEVPTDFTVHSLP--ETTUVYPADVVFEG 139
DB 156 RKGF--PESDYRRLTFVADIKGKQEVAPVYSHLYIDVPQKLTVRP-DILIVEG 212
QY 140 ILVFSQ-----EIRDMFLRLFVDT-DSDVR---LSR-RVLR----- 172
DB 213 LNVLPALPGKDGRTVGLADYFDFSVYVDARPEIDRWYLNFRKLRATAFQNPSSYFR 272
QY 173 ---DYRGRDLQTLTQYTFVKRPFEECLPTKKYADVILPRVDMMV 218
DB 273 RYGVSEDAEDVARTWRTINKVLLDENVAFTGRALVVRKGPDKV 321
```

RESULT 9

```
US-09-734-569-88
; Sequence 88, Application US/09734569
; Patent No. US20020064816A1
; GENERAL INFORMATION:
; APPLICANT: Lerchl, Jens
; APPLICANT: Renz, Andreas
; APPLICANT: Ehrhardt, Thomas
; APPLICANT: Reindl, Andreas
; APPLICANT: Citpus, Petra
; APPLICANT: Bischoff, Friedrich
; APPLICANT: Frank, Markus
; APPLICANT: Freund, Annette
; APPLICANT: Duwenig, Elke
; APPLICANT: Schmidt, Ralf-Michael
; APPLICANT: Reeki, Ralf
; TITLE OF INVENTION: Moss genes from Physcomitrella patens encoding proteins involved
; FILE REFERENCE: BASF-NAE-1332-99-US
; CURRENT APPLICATION NUMBER: US/09/734, 569
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 60/171,101
; PRIOR FILING DATE: 1999-12-16
; NUMBER OF SEQ ID NOS: 181
; SOFTWARE: PatentIn Ver. 2.1/WordPerfect 6.1
```

```
; SEQ ID NO 88
; LENGTH: 143
; TYPE: PRT
; ORGANISM: Physcomitrella patens
US-09-734-569-88
```

```
Query Match 8.2%; Score 119.5; DB 9; Length 143;
Best Local Similarity 26.8%; Pred. No. 6.4e-05;
Matches 37; Conservative 28; Mismatches 52; Indels 21; Gaps 5;
```

```
QY 25 LIGVSGGTASGKSTVCEKIMELG-----QNEVEQRQRKVVLSQDFYKVL 72
DB 14 VIGLADSGGKSTFMRRLTSVCGATPCKGNPDSNTLISPTTVICL--DYSLSLR 71
QY 73 AEOKAKLKGQVNFDPDAFNDMDHRTLNKIVEGKTVVEVPTDFTVH-SRLPETTVVP 131
DB 72 YGRKEKAVNAL-----DPRANFDMVEQVKALKEGKSEKPYNHVTGLDAPET--IHP 125
QY 132 ADVVLEFEGILVFYSQEIR 149
DB 126 PKLVIEGLHPMYDERVR 143
```

RESULT 10

```
US-10-301-822-217
; Sequence 217, Application US/10301822
; Publication No. US20030148410A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Berger, Allison
; APPLICANT: Guillemette, Tracy L.
; APPLICANT: Kamackar, Shubhangi
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Thibodeau, Stephen N.
; APPLICANT: Burgart, Lawrence J.
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF COLON CANCER
; FILE REFERENCE: MPW01-029P2RM
; CURRENT APPLICATION NUMBER: US/10/301, 822
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 60/339, 971
; PRIOR FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: US 60/361, 978
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/381, 988
; PRIOR FILING DATE: 2002-05-20
; NUMBER OF SEQ ID NOS: 228
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 217
; LENGTH: 216
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-301-822-217
```

```
Query Match 6.7%; Score 97; DB 12; Length 216;
Best Local Similarity 24.4%; Pred. No. 0.036;
Matches 53; Conservative 36; Mismatches 64; Indels 64; Gaps 12;
```

```
QY 2 ASAGE-DESPAPEDRHPQRFLLGVSGGTASGKSTVCEKIMELGQ-----NEVEQR 55
DB 36 ASAGPERDARPGAFAL-----GTWGAHSASEVREBLEGKTGRSSDOIQL 81
QY 56 QRKVVLSQDFYKVLTAEOKAKALKGQVNFDPDA-----FNDMDHRTL 101
DB 82 HRR-----FKQLSGDQPTIR---KENFNNVPDLLENPFRSKTVRAFQDN---RNL 125
QY 102 KNIVEGKTVVEVPTDFTVH-SRLPETTVVPADVLEFEGILVFYSQEIRDMFHLRLFV 158
DB 126 RKGPSGLADEINDEFLLTMSYFRPIDTWDR-----EYVELSRKEXLRFLFHM---Y 175
QY 159 DTDSDVRLSRVLRDVRGRD-LEQILTQYTFVKPA 194
```


Db 176 DSDSDGRITLSEYRNVMKMSRSCCETLTSTRSLAPFS 212

RESULT 11
US-10-301-822-220

; Sequence 220, Application US/10301822
; Publication No. US20030148410A1

; GENERAL INFORMATION:

; APPLICANT: Millennium Pharmaceuticals, Inc.

; APPLICANT: Berger, Allison

; APPLICANT: Guillemette, Tracy L.

; APPLICANT: Kamatkar, Shubhangi

; APPLICANT: Schlegel, Robert

; APPLICANT: Monahan, John E.

; APPLICANT: Thibodeau, Stephen N.

; APPLICANT: Burgart, Lawrence J.

; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND

; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND

; FILE REFERENCE: MEM01-029P2RNM

; CURRENT APPLICATION NUMBER: US/10/301,822

; CURRENT FILING DATE: 2002-11-21

; PRIOR APPLICATION NUMBER: US 60/339,971

; PRIOR FILING DATE: 2001-12-10

; PRIOR APPLICATION NUMBER: US 60/361,978

; PRIOR FILING DATE: 2002-03-05

; PRIOR APPLICATION NUMBER: US 60/381,988

; PRIOR FILING DATE: 2002-05-20

; NUMBER OF SEQ ID NOS: 228

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 220

; LENGTH: 251

; TYPE: PRT

; ORGANISM: Homo Sapiens

US-10-301-822-220

Query Match 6.6%; Score 95; DB 12; Length 251;

Best Local Similarity 24.6%; Pred. No. 0.075; Matches 51; Conservative 35; Mismatches 53; Indels 68; Gaps 12;

QY 2 ASAGGE-DCESPAEADRPQRFLIGVSGTASGKSTVCEKIMELGQ-----NEVEOR 55

Db 36 ASAGPERDARPGAPAA-----GTMAAHSASEBEVRELEKGTGSSDQIEQL 81

QY 56 QKRVILISDRFYKVLTAEOKAKALKGQYNFDH-PDA-----FNDLMHRTL 101

Db 82 HRR-----FKQLSGDQPTIR---KENFNVPDLNPIRSKIVRAFFDN---RNL 125

QY 102 KNIVEGKTVETPTVDVT---HSRLPETTVVPADVLFEGILVFSQETIRDMFHLRLFV 158

Db 126 RKGPSSGLADEINFEDELITMSYFRPIDTWD-----EVELSRKEXLRFLFHM---Y 175

QY 159 DTDSDVRLSRVLDVRRGRDLEQILT 185

Db 176 DSDSDGRITLSEYRNVM-----VEELLS 197

RESULT 12
US-10-106-698-6128

; Sequence 6128, Application US/10106698

; Publication No. US20030109690A1

; GENERAL INFORMATION:

; APPLICANT: Ruben et al.

; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide

; FILE REFERENCE: PA0005P1

; CURRENT APPLICATION NUMBER: US/10/106,698

; CURRENT FILING DATE: 2002-03-27

; PRIOR APPLICATION NUMBER: PCT/US00/26524

; PRIOR FILING DATE: 2000-09-28

; PRIOR APPLICATION NUMBER: US 60/157,137

; PRIOR FILING DATE: 1999-09-29

; PRIOR APPLICATION NUMBER: US 60/163,280

; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0

; SEQ ID NO 6128

; LENGTH: 257

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-106-698-6128

Query Match 6.6%; Score 95; DB 15; Length 257;

Best Local Similarity 24.6%; Pred. No. 0.077; Matches 51; Conservative 35; Mismatches 53; Indels 68; Gaps 12;

QY 2 ASAGGE-DCESPAEADRPQRFLIGVSGTASGKSTVCEKIMELGQ-----NEVEOR 55

Db 26 ASAGPERDARPGAPAA-----GTWGAHSASEBEVRELEKGTGSSDQIEQL 71

QY 56 QKRVILISDRFYKVLTAEOKAKALKGQYNFDH-PDA-----FNDLMHRTL 101

Db 72 HRR-----FKQLSGDQPTIR---KENFNVPDLNPIRSKIVRAFFDN---RNL 115

QY 102 KNIVEGKTVETPTVDVT---HSRLPETTVVPADVLFEGILVFSQETIRDMFHLRLFV 158

Db 116 RKGPSSGLADEINFEDELITMSYFRPIDTWD-----EVELSRKEXLRFLFHM---Y 165

QY 159 DTDSDVRLSRVLDVRRGRDLEQILT 185

Db 166 DSDSDGRITLSEYRNVM-----VEELLS 187

RESULT 13

US-10-301-822-215

; Sequence 215, Application US/10301822

; Publication No. US20030148410A1

; GENERAL INFORMATION:

; APPLICANT: Millennium Pharmaceuticals, Inc.

; APPLICANT: Berger, Allison

; APPLICANT: Guillemette, Tracy L.

; APPLICANT: Kamatkar, Shubhangi

; APPLICANT: Schlegel, Robert

; APPLICANT: Monahan, John E.

; APPLICANT: Thibodeau, Stephen N.

; APPLICANT: Burgart, Lawrence J.

; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND

; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND

; FILE REFERENCE: MEM01-029P2RNM

; CURRENT APPLICATION NUMBER: US/10/301,822

; CURRENT FILING DATE: 2002-11-21

; PRIOR APPLICATION NUMBER: US 60/339,971

; PRIOR FILING DATE: 2001-12-10

; PRIOR APPLICATION NUMBER: US 60/361,978

; PRIOR FILING DATE: 2002-03-05

; PRIOR APPLICATION NUMBER: US 60/381,988

; PRIOR FILING DATE: 2002-05-20

; NUMBER OF SEQ ID NOS: 228

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 215

; LENGTH: 267

; TYPE: PRT

; ORGANISM: Homo Sapiens

US-10-301-822-215

Query Match 6.6%; Score 95; DB 12; Length 267;

Best Local Similarity 24.6%; Pred. No. 0.082; Matches 51; Conservative 35; Mismatches 53; Indels 68; Gaps 12;

QY 2 ASAGGE-DCESPAEADRPQRFLIGVSGTASGKSTVCEKIMELGQ-----NEVEOR 55

Db 36 ASAGPERDARPGAPAA-----GTWGAHSASEBEVRELEKGTGSSDQIEQL 81

QY 56 QKRVILISDRFYKVLTAEOKAKALKGQYNFDH-PDA-----FNDLMHRTL 101

Db 82 HRR-----FKQLSGDQPTIR---KENFNVDLELNPISRKIVAFEDN---FNL 125
Qy 102 KNIVEGKTVETVYQFVT---HSRLPETTVYVPADVLFEGILVYSQEIFDMFLRLFV 158
Db 126 RKGPGGLADEINFEEPLTMSYFRPIDYTMDE-----EQLVSRKEKLFLEFMM---Y 175
Qy 159 DTSDVRLSRVLRDVRGRDLQILT 185
Db 176 DSDSDGRITLBEYRNV-----VEELIS 197

RESULT 14

US-09-769-744A-64
; Sequence 64, Application US/09769744A
; Publication No. US20030134407A1
; GENERAL INFORMATION:

; APPLICANT: Le Page, Richard WF
; APPLICANT: Wells, Jeremy M
; APPLICANT: Hamblify, Sean B
; APPLICANT: Hambro, Philip M
; TITLE OF INVENTION: Proteins
; FILE REFERENCE: PWC/P21122MO
; CURRENT APPLICATION NUMBER: US/09/769,744A
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: PCT/GB99/02452
; PRIOR FILING DATE: 1999-07-27
; PRIOR APPLICATION NUMBER: GB 9816336.3
; PRIOR FILING DATE: 1998-07-27
; PRIOR APPLICATION NUMBER: US 60/125329
; PRIOR FILING DATE: 1999-03-19
; NUMBER OF SEQ ID NOS: 196
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 64
; LENGTH: 210
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
; US-09-769-744A-64

Query Match 6.4%; Score 93; DB 12; Length 210;

Best Local Similarity 21.0%; Pred. No. 0.095; 79; Indels 66; Gaps 11;

Matches 49; Conservative 39; Mismatches 79; Indels 66; Gaps 11;

Qy 25 LIGVSGTSGKSTVCEKIMELQNEVEQORQKVILISOD-----REFKYVTA 73
Db 13 IIGITGGIAGSGKSTVYNFL-----RQSGFQVVDADAVVHQLQKPGGRLEFALVQ 61
Qy 74 E--QKAKALKGQYN-----FDHPDAFNDLMHRTLAKIVGKTV--EVPYTFVTHS 121
Db 62 HFGQELIENGEILNRPILASLIFSND-----EREWSKOIGSEILIREELAT---LRE 110
Qy 122 RLPEFTVYVPADVLFEGILVYSQEIFDMFLRLFVDTSDVRLSRVLRDVRGRDL 181
Db 111 QLAQTEIEFFMDIPL-----LFEDYSDYDMFAETMIVYVDDAQVERLMKRDQSLKDEAE 164
Qy 182 -QILTQYTFVYKPAEEFCLEPTKYADVIRPGVDNMVAIILVQHIDILNG 233
Db 165 SRLAQQW-----PLEKKD-LASQVLDNNGNQQLNQVHILLEG 203

RESULT 15

US-09-815-242-5691

; Sequence 5691, Application US/09815242

; Patent No. US20020061569A1

; GENERAL INFORMATION:

; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in

; TITLE OF INVENTION: Prokaryotes

; FILE REFERENCE: ELITRA.011A

; CURRENT APPLICATION NUMBER: US/09/815,242

; CURRENT FILING DATE: 2001-03-21

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/242,578

; PRIOR FILING DATE: 2000-10-23

; PRIOR APPLICATION NUMBER: 60/253,625

; PRIOR FILING DATE: 2000-11-27

; PRIOR APPLICATION NUMBER: 60/257,931

; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 60/269,308

; PRIOR FILING DATE: 2001-02-16

; NUMBER OF SEQ ID NOS: 1410

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 5691

; LENGTH: 433

; TYPE: PRT

; ORGANISM: Staphylococcus aureus

US-09-815-242-5691

Query Match 6.4%; Score 93; DB 9; Length 433;

Best Local Similarity 23.8%; Pred. No. 0.28; 93; Indels 42; Gaps 9;

Matches 53; Conservative 35; Mismatches 93; Indels 42; Gaps 9;

Qy 31 GTASGKSTVCEKIMELQNEVEQORQKVILISQDRFYKVLTAQKAKALKGQYNFDPD 90
Db 174 GSTHSNSTWCKRSTLAEQLKI---QLFV-----ESFLHAFETNSSL-----YVHP 218
Qy 91 APNDLMHRTLAKIVGKTVETVYQFVTSHRLPET--TVVYVPADVLFEGILVYSQEIFR 149
Db 219 LFMNDF---SLKAFPEG--TDVPVYVYKLFPEEGPITMTILIREWRLMWKEMAILQAFRV 273
Qy 150 DMFHLRLFVDTSDVRLSRVLRDVRGRDLBQILQYTFVYKPA----- 194
Db 274 SVNLQPMVWENPVRPBTDEGDIEHFEILPILQELVLYVAYYTAIILDPFSQDPENGH 333
Qy 195 -PEEFCLPTKKY---ADVI-IPRGVDNMVAIILVQHIDIL 231
Db 334 YFDFSAVPFQYVKNQGDVVQIPRMSSEDYRTAMIGHIKWML 376

Search completed: November 25, 2003, 08:08:11

Job time : 33 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OW protein - protein search, using sw model

Run on: November 25, 2003, 08:01:09 ; Search time 20 Seconds
(without alignments)
586.005 Million cell updates/sec

Title: US-09-896-522-2

Perfect score: 1450

Sequence: 1 MASAGEDCESPAPADRP...HPGMLTSGKSHLESSRPH 277

Scoring table: BLOSUM62

Searched: Gapop 10.0 , Gapext 0.5

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

- Issued Patents NA:*
- 1: /cgn2_6/prodata/2/iaa/5A COMB.pep:*
 - 2: /cgn2_6/prodata/2/iaa/5B COMB.pep:*
 - 3: /cgn2_6/prodata/2/iaa/6A COMB.pep:*
 - 4: /cgn2_6/prodata/2/iaa/6B COMB.pep:*
 - 5: /cgn2_6/prodata/2/iaa/PCTUS COMB.pep:*
 - 6: /cgn2_6/prodata/2/iaa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1363	94.0	260	4 US-09-536-647-2	Sequence 2, App11
2	1325	91.4	277	4 US-09-536-647-3	Sequence 3, App11
3	451	31.1	216	4 US-09-134-001C-3618	Sequence 3618, Ap
4	434	29.9	210	4 US-09-107-532A-5024	Sequence 5024, Ap
5	412	28.4	222	4 US-09-198-452A-7835	Sequence 783, App
6	99.5	6.9	208	4 US-09-134-001C-3785	Sequence 3785, Ap
7	93.5	6.4	216	4 US-09-252-991A-26519	Sequence 26519, A
8	91.5	6.3	1376	1 US-08-420-235B-3	Sequence 3, App11
9	91.5	6.3	1376	3 US-08-793-624-3	Sequence 3, App11
10	91.5	6.3	1376	5 PCT-US95-10194-3	Sequence 3, App11
11	89	6.1	517	4 US-09-252-991A-22249	Sequence 22249, A
12	88	6.1	674	3 US-08-961-083-200	Sequence 200, App
13	88	6.1	674	4 US-09-536-784-200	Sequence 200, App
14	86	5.9	542	4 US-09-198-452A-1003	Sequence 1003, Ap
15	85	5.9	802	4 US-09-134-001C-3840	Sequence 3840, Ap
16	85	5.9	810	4 US-09-540-824-25	Sequence 25, App1
17	84.5	5.8	647	4 US-09-134-001C-5458	Sequence 5458, Ap
18	84.5	5.8	2639	3 US-09-080-983-3	Sequence 3, App11
19	84	5.8	687	2 US-08-540-804-4	Sequence 3, App11
20	84	5.8	687	2 US-08-218-265-4	Sequence 4, App11
21	84	5.8	687	3 US-08-521-872-4	Sequence 4, App11
22	84	5.8	687	3 US-08-590-399-4	Sequence 4, App11
23	81	5.6	937	4 US-09-107-532A-4335	Sequence 4335, Ap
24	80.5	5.6	625	4 US-09-252-991A-20865	Sequence 20865, A
25	80.5	5.6	661	4 US-09-328-352-5666	Sequence 5666, Ap
26	80.5	5.6	722	4 US-08-817-832B-32	Sequence 32, App1
27	80	5.5	205	4 US-09-328-352-8081	Sequence 8081, Ap

28	78.5	5.4	284	4 US-08-858-207A-346	Sequence 346, App
29	78.5	5.4	285	4 US-09-328-352-4464	Sequence 4464, Ap
30	78.5	5.4	414	4 US-09-647-540A-6	Sequence 6, App11
31	78.5	5.4	414	4 US-10-119-600-6	Sequence 6, App11
32	78.5	5.4	607	3 US-08-781-891-75	Sequence 75, App1
33	78.5	5.4	607	3 US-09-618-166-75	Sequence 75, App1
34	78.5	5.4	722	4 US-09-984-890-4	Sequence 4, App11
35	78.5	5.4	724	4 US-09-984-890-2	Sequence 2, App11
36	78.5	5.4	745	4 US-09-523-849-36	Sequence 36, App1
37	78.5	5.4	779	1 US-08-426-627-4	Sequence 4, App11
38	78.5	5.4	779	1 US-08-426-627-24	Sequence 24, App11
39	78.5	5.4	836	1 US-08-426-627-6	Sequence 6, App11
40	78.5	5.4	837	1 US-08-426-627-23	Sequence 23, App1
41	78.5	5.4	2482	1 US-08-328-254-6	Sequence 6, App11
42	78.5	5.4	3248	1 US-08-353-700-1	Sequence 1, App11
43	78.5	5.4	3248	5 PCT-US95-16216-1	Sequence 1, App11
44	78	5.4	237	3 US-09-320-878-11	Sequence 11, App1
45	78	5.4	237	3 US-09-105-537-18	Sequence 18, App1

ALIGNMENTS

RESULT 1
US-09-536-647-2
; Sequence 2, Application US/09536647
; Patent No. 6579708
; GENERAL INFORMATION:
; APPLICANT: Ho, Yen Sen
; APPLICANT: Johnson, Randall
; TITLE OF INVENTION: No. 6579708el Human Uridine Kinase
; FILE REFERENCE: GP50020
; CURRENT APPLICATION NUMBER: US/09/536,647
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 260
; TYPE: PRT
; ORGANISM: Human
US-09-536-647-2

Query Match 94.0%; Score 1363; DB 4; Length 260;
Best Local Similarity 100.0%; Pred. No. 1e-150;
Matches 260; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MASAGEDCESPAPADRP...PFLIGVSGTASGKSTVCCKIMELLGQNEVEQRQKVV	60
DB	1	MASAGEDCESPAPADRP...PFLIGVSGTASGKSTVCCKIMELLGQNEVEQRQKVV	60
QY	61	IIISDRFYKVLTAEOAKALKGOYNFDHPDANDNDIMRTLNKIVHGKTVVPTVDFVTH	120
DB	61	IIISDRFYKVLTAEOAKALKGOYNFDHPDANDNDIMRTLNKIVHGKTVVPTVDFVTH	120
QY	121	SRLEPTTVYPPADVILFESGILVFYSOETIRDMFHLRLFVTDSDVRLSRVLDVVRGRDL	180
DB	121	SRLEPTTVYPPADVILFESGILVFYSOETIRDMFHLRLFVTDSDVRLSRVLDVVRGRDL	180
QY	181	EQILQYTTFFVAPAEFECLPTKRYADVILIPRGVDMVAINILVQHIQDILNGDICKMHR	240
DB	181	EQILQYTTFFVAPAEFECLPTKRYADVILIPRGVDMVAINILVQHIQDILNGDICKMHR	240
QY	241	GGSNGRSYKRTSPSRQDHPG	260
DB	241	GGSNGRSYKRTSPSRQDHPG	260

RESULT 2
US-09-536-647-3
; Sequence 3, Application US/09536647
; Patent No. 6579708
; GENERAL INFORMATION:
; APPLICANT: Ho, Yen Sen

```

; APPLICANT: Johnson, Randall
; TITLE OF INVENTION: No. 6579708e1 Human Uridine Kinase
; FILE REFERENCE: GP50020
; CURRENT APPLICATION NUMBER: US/09/536,647
; CURRENT FILING DATE: 2000-03-27
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 277
; TYPE: PRT
; ORGANISM: Human
US-09-536-647-3

Query Match          91.4%; Score 1325; DB 4; Length 277;
Best Local Similarity 91.7%; Pred. No. 3e-146;
Matches 254; Conservative 11; Mismatches 12; Indels 0; Gaps 0;

QY 1 MASAGDECSAPAEADRPQRPLIGVSGTASGKTVCEKIMELIGONEVEORQRKV 60
DB 1 MASAGGGSSEAPAEADRPQRPPLIGVSGTASGKTVCEKIMELIGONEVEORQKLV 60
QY 61 ILSQDREYKVLTAQKAKALKQGYNPDHPDAFNDMLMRTLKNIVEGKTVETPTDEVT 120
DB 61 ILSQDCFYKVLTAQKAKALKQGYNPDHPDAFNDMLMRTLKNIVEGKTVETPTDEVT 120
QY 121 SRLPETHVVPADVVLREGILVFYSQERDMFHLRLPVDITSDVRLSRVLRDVRGRDL 180
DB 121 SRLPETHVVPADVVLREGILVFYSQERDMFHLRLPVDITSDVRLSRVLRDVRGRDL 180
QY 181 EQLITQYTFVYKPAFEFECLEPTKKYADVIIPRGVNNVAINLYQHIQDILNGDICKMHR 240
DB 181 EQLITQYTFVYKPAFEFECLEPTKKYADVIIRGVDNNVAINLYQHIQDILNGDICKRHR 240
QY 241 GGSNGRSYKRTSEPGDHGMLTSGKSHLESSSRPH 277
DB 241 GGPNGRNHKKRTFPEPGDHPGLATGKSHLESSSRPH 277

RESULT 3
US-09-134-001C-3618
; Sequence 3618, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3618
; LENGTH: 216
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3618

Query Match          31.1%; Score 451; DB 4; Length 216;
Best Local Similarity 42.9%; Pred. No. 2.4e-44;
Matches 91; Conservative 44; Mismatches 61; Indels 16; Gaps 4;

QY 25 LIGVSGGTASGKTVCEKIMELIGONEVEORQRKVILSODRFYK--VITAEQKAKALK 81
DB 15 IIGAGSGSGSKTVTAIMKNL-----EGHSVALLAODYKKDSHLSFEERLET-- 65
QY 82 GQYVFDHPDAFNDMLMRTLKNIVEGKTVETPTDPTVTHSRLPETHVVPADVVLREGIL 141
DB 66 ---NYDHPFAFDNDLLIHNMLDLNRNGKPEVETPDYDSQHSKETIIFDPPDVIVIGEIF 122
QY 142 VFYSQERDMFHLRLPVDITSDVRLSRVLRDVR-RGRDLQILITQYTFVYKPAFEFCL 200
```

```

DB 123 ALNNTRLDMMDVKIYVDDADARILRLTRDKEGRMTSEYINQYLVNVRMHEQFLE 182
QY 201 PTKKADVIIIPRGVNNVAINLYQHIQDILN 232
DB 183 PTKKHADIIIPGGSNKNVAIDIMTKIOSIVS 214

RESULT 4
US-09-107-532A-5024
; Sequence 5024, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Arinello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 5024:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 210 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (8) LOCATION 1...210
; SEQUENCE DESCRIPTION: SEQ ID NO: 5024:
US-09-107-532A-5024

Query Match          29.9%; Score 434; DB 4; Length 210;
Best Local Similarity 38.6%; Pred. No. 2.2e-42;
Matches 83; Conservative 51; Mismatches 65; Indels 16; Gaps 4;

QY 22 RPLIGVSGGTASGKTVCEKIMELIGONEVEORQRKVILSODRFYK--VITAEQKAK 78
DB 6 KPIIIGVTGSGSGKTSVSRALFNPDPH-----SIMMLEQDSYKKDSHLSFEERLNL 58
QY 79 ALKGQYNPDHPDAFNDMLMRTLKNIVEGKTVETPTDPTVTHSRLPETHVVPADVVLFE 138
DB 59 T-----NYDHPFAFDNDLLIQHVGDLNLNKAIEKPYDVVAHTRSQATITIQEKEVITLE 113
QY 139 GILVFYSQERDMFHLRLPVDITSDVRLSRVLRDVR-RGRDLQILITQYTFVYKPAFEF 197
```

Db 114 GILIEDELRLMDIKYVDDIDIRIRIKRMERKGLDSVIEQYITLVKPMYHQ 173
Db 174 FIEPTKRYADIVPEGGENHVAIDLITTKVASFNL 208
Dy 198 FCLPTKKYADVIIIPRGVDMMALNLVQHIDILN 232
 | | | : | | : | : | | | : | |
 | | | : | | : | : | | | : | |

```

RESULT 5
US-09-198-452A-783
: Sequence 783, Application US/09198452A
: Patent No. 6559294
: GENERAL INFORMATION:
: APPLICANT: Griffois, R.
: TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
: TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention
: TITLE OF INVENTION: and treatment of infection
: FILE REFERENCE: 9710-003-999
: CURRENT APPLICATION NUMBER: US/09/198,452A
: CURRENT FILING DATE: 1998-11-24
: NUMBER OF SEQ ID NOS: 6849
: SEQ ID NO 783
: LENGTH: 222
: TYPE: PR1
: ORGANISM: Chlamydia pneumoniae
US-09-198-452A-783

```

Query Match	28.4%;	Score 412;	DB 4;	Length 222;
Best Local Similarity	38.4%;	Pred. No. 9.1e-40;		
Matches 84; Conservative	54;	Mismatches 59;	Indels 22;	Gaps 6

```

QY 25 LIIVSGGIAAGKSTVECKTEKIMELLGQNEBQROKVVILISODEFYK---VLTQEQCAKKK 81
Db 8 IITGTGSGAGKTTLTQNIKEIFGD-----VSVIQDNYKRRSHYTPPEBANLI 58
QY 82 GQVNEHPRAFDDMDHRLTKNLVEKTYVEPPEYTHSR-LPEPTVVPADVILPEFGI 144
Db 59 ----WHPRAFDDMDLISDKIKLKNNEIYQAFPEFDPVLGNRSKTELETTSYVILVGGI 111

```

Dy 141 LVFYSQELRDMFLRLFPVDTSDVRSRVLRDVR-RGRDLQLITQYTFPKPAFEFC 199

Dz 115 LVEENQELRDLMDIRIFVDTDADERILRRWRDVQEGGDSYDCIMSRYLSTWKPMEKEFI 174

```

QY      200 LPTKKYADVII PRGVDNVAI NLIVQ---HIQDILNGD 234
      ||:||||:| :|:| :|:|
Db      175 EPTRYADIIVHGNYRQNVVTNII LSKIKNHLLENALIESD 213

```

```

RESULT 6
US-09-134-001C-3785
; Sequence 3785, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; PENDING FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3785
; LENGTH: 208
; TYPE: prt
; ORGANISM: Staphylococcus epidermidis
; US-09-134-001C-3785

```

Query Match	6.9%;	Score 99.5;	DB 4;	Length 208;
Best Local Similarity	21.5%;	Pred. No. 0.0026;		
Matches 40; Conservative	48;	Mismatches 71;	Indels 27;	Gaps 8;

```
QY      25 LIGVSGGNAKSKWJCE-----XIMEL-LGONEVEQRQRKVLLSQDRFVKVLTAE-- 74
        :|::|||:|||::||::||::||::||::||::||::||::||::||::||::||
Db      9 VIGITGGTATGSIVSEBELLTAYGFKEIVPADIASRBAVKKSGLQEVEKI FEEBALDENG 68
        :|::|||:|||::||::||::||::||::||::||::||::||::||::||
QY      75 QKAVALKQGVNFHDHDAFD--NDLMHRTLKNIVEEK-----TYEVP-TYDFVTN 120
        :|::|||:|||::||::||::||::||::||::||::||::||::||
Db      69 EMNQGYVEIYFNPHDLREALNEIHPIVRELMEQEKNNYLHGHWINDIPLYENELQ 128
        :|::|||:|||::||::||::||::||::||::||::||::||::||
QY      121 SRLPETYVYPADVULFEGLVFSQELRDMEHLFLF--VDTSDYRLSRRLRVDRYGR 178
        :|::|||:|||::||::||::||::||::||::||::||::||::||
Db      129 DTVEVWWVVVYSBSIQIDRLMERNNLSLEDA-KARVVSQISIDKKSRMADHVINDLGDKL 187
        :|::|||:|||::||::||::||::||::||::||::||::||::||
QY      179 DLEQLL 184
        :|::||
Db      188 ELKQNL 193
```

```

RESULT 7
US-09-252-991A-26519
; Sequence 26519, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIORITY FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIORITY FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIORITY FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 26519
; LENGTH: 216
; TYPE: PRF
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-26519

```

Query Match	6.4%;	Score 93.5;	DB 4;	Length 216;
Best Local Similarity	23.7%;	Pred. No. 0.014;		
Matches 41;	Conservative 31;	Mismatches 74;	Indels 27;	Gaps 6;

```
Qy      22 RPELLIGVSGSTASGKSTVECKIMEL--LQONEVEQOROKVVLISODRFYV-----L    71
          : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      16 QPWLGLTGIGSGKSAAAEHFISSLGVHLVDADHDARWVEGRPALAKIVERFGDILL    75
          : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy      72 TAQQAALKAGQYNFPHPRDAFNDMLNHTLKIVVEKTYEVEPTDFVTHSRLEPTTVVP   13
          : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db       76 PDGQLDRALRERIFQAPF--ERRMVEOLLHPILGAETIQ----YLAAESPVALIVSP  12
          : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy     132 ADVLVLEGILVFYSOEIRDMFHRLFPVDPDSVRISRRLRVLRVGRRDEIOL    184
          : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db     129 -----LLVESGQ---RQMTWRVLVVDTPEHIQDLQRITLKRKVSEBOVRSLT    171
```

RESULT 8
 US-08-420-235B-3
 Sequence 3, Application US/08420235B
 Patent No. 5801042
 GENERAL INFORMATION:
 APPLICANT: Yuan
 APPLICANT: Chang, Patrick S.
 TITLE OF INVENTION: UNIQUE ASSOCIATED KAPOSI'S SARCOMA VIRUS
 TITLE OF INVENTION: SEQUENCES AND USES THEREOF
 NUMBER OF SEQUENCES: 47
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Cooper & Dunham LLP
 STREET: 1185 Avenue of the Americas
 CITY: New York
 STATE: New York
 COUNTRY: U.S.A.
 ZIP: 10036

[illegible]

```

RESULT 11
US-09-252-991A-22249
; Sequence 22249, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS.
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS.
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 22249
; LENGTH: 517
; TYPE: PRF
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-22249

```

Query March	6.1%;	Score 89;	DB 4;	Length 517;
Best Local Similarity	23.3%;	Pred. No. 0.18;		
Matches	59;	Conservative	30;	Mismatches 102;
			Indels	62;
			Gaps	11

QY	32	TASGKSTVCEKIMELLGONEVEQR----	QRRVLLSDRFYKVLTAERKAKALGXQYN-	85
DB	56	TAAISQALKRIKIEIEPSEARQNFHPQLRIGYRCGEVSTTRAYAKVQ-VGEYSV		114
QY	86	-FDHPDAPDNLMHRTLKNIYEGKIVEPYTDVTHSLRPETTVYYPADVILFEGILV--		142
DB	115	TVTHPDAFRKYLLEQ-LRPLMDDFVQVQGP--SQODI PYPYVEGGELAGSGVTAAE		171
QY	143	-----FYSQIRIMFHLRFVDPTSDVRLSRRLVDVRRGRDLEQ		182
DB	172	LARVPSTDSLAAQDIDLADGLYEMERAOQLPLALDARVPSLRVLHYTGSDMRWQYP		231
QY	183	--ILTYTTPV-----KPAEEFCFLPKTKADVILPGVDMMVAINILIVQHI		227
DB	232	WILLTNHRYVDQFRLGLTRLREDPRFVRWLP---GNVILIERGMDGEQA-NAIVAGV		286
QY	228	QDILNGDICKWHR		240
DB	287	-----VMHR		290

RESULT 12
US-08-961-083-200
Sequence 200, Application US/08961083
Patent No. 6159469
GENERAL INFORMATION:
APPLICANT: Choi et. al.
TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
NUMBER OF SEQUENCES: 452
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland

```

?      COUNTRY:  USA
?      ZIP:      20850
?      COMPUTER READABLE FORM:
?      MEDIUM TYPE:  Diskette, 3.50 inch, 1.4Mb storage
?      COMPUTER:  HP Vectra 486/33
?      OPERATING SYSTEM:  MSDOS version 6.2
?      SOFTWARE:  ASCII Text
?      CURRENT APPLICATION DATA:
?      APPLICATION NUMBER:  US/08/961,083
?      FILING DATE:
?      CLASSIFICATION:  435
?      PRIOR APPLICATION DATA:
?      APPLICATION NUMBER:
?      FILING DATE:
?      ATTORNEY/AGENT INFORMATION:
?      NAME:  Brookes, A. Anders
?      REGISTRATION NUMBER:  36,373
?      REFERENCE/DOCKET NUMBER:  PB340P2
?      TELECOMMUNICATION INFORMATION:
?      TELEPHONE:  (301) 309-8504
?      TELEFAX:  (301) 309-8512
?      INFORMATION FOR SEQ ID NO:  200:
?      SEQUENCE CHARACTERISTICS:
?      LENGTH:  674 amino acids
?      TYPE:  amino acid
?      STRANDEDNESS:  single
?      TOPOLOGY:  linear
?      MOLECULE TYPE:  protein
?      US-08-961-083-200

```

```

Query Match          6.1%; Score 88; DB 3; Length 674;
Best Local Similarity 27.5%; Pred. No. 0.36;
Matches 30; Conservative 16; Mismatches 35; Indels 28; Gaps 6

Cy 105 VEGTVEVPTDYD-----VTHSLPETTVYPADVVLFSGILVYFSOEIRDMFHLRLFV 158
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 562 IETLVOSIPDQFIARAVIVSDRLAQWT-IYP-NVTIFNGI-----HYLV 605

Cy 159 DTSDVSLSRVLDVVRGRDLDEILQYTTFFVP--AFEEFCUPTKXY 205
      | : : : : : : : : : : : : : : : : : : : : : : : :
Db 606 DVNDLVEVTSQVLLDINHGEKTEBILQFANLGRKPLISFEN---TKTY 650

RESULT 13
US-09-536-784-200
; Sequence 200, Application US/09536784
; Patent No. 6573082
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigenes and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/536,784
; FILING DATE: 30-Oct-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/961,083
; FILING DATE: OCT-30-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Michelle S. Marks
; REGISTRATION NUMBER: 41,971
;

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2003, 07:54:09 ; Search time 40 Seconds
(without alignments)
1099.182 Million cell updates/sec

Title: US-09-896-522-2
Perfect score: 1450
Sequence: 1 MASAGGEDESPAPADRP.....HPGMLTSGKSHLESSSRPH 277

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A_Geneseq_19Jun03:*

- 1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
- 2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
- 3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
- 4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
- 5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
- 6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
- 7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
- 8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
- 9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
- 10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
- 11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
- 12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
- 13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
- 14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
- 15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
- 16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
- 17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
- 18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
- 19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
- 20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
- 21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
- 22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
- 23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*
- 24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1450	100.0	277	22	AA19502 Human polypeptide
2	1450	100.0	277	22	AA19502 Human polypeptide
3	1450	100.0	277	22	AA19502 Human polypeptide
4	1450	100.0	277	22	AA19502 Human polypeptide
5	1372.5	94.0	276	22	AA19502 Human polypeptide
6	1363	91.4	276	22	AA19502 Human polypeptide
7	1325	91.4	276	22	AA19502 Human polypeptide
8	941	64.9	190	23	AA19502 Human polypeptide
9	934	64.4	261	22	AA19502 Human polypeptide

10	934	64.4	261	22	AA19502 Human polypeptide
11	901	62.1	337	23	AA19502 Human polypeptide
12	896	61.8	337	22	AA19502 Human polypeptide
13	889	61.3	335	22	AA19502 Human polypeptide
14	738	50.9	326	22	AA19502 Human polypeptide
15	589	40.6	326	22	AA19502 Human polypeptide
16	579	39.9	120	23	AA19502 Human polypeptide
17	578	39.9	481	21	AA19502 Human polypeptide
18	578	39.9	490	21	AA19502 Human polypeptide
19	578	39.9	512	21	AA19502 Human polypeptide
20	570.5	39.3	500	21	AA19502 Human polypeptide
21	570.5	39.3	511	21	AA19502 Human polypeptide
22	545.5	37.6	548	23	AA19502 Human polypeptide
23	542	37.4	614	22	AA19502 Human polypeptide
24	542	37.4	614	22	AA19502 Human polypeptide
25	540.5	37.3	548	23	AA19502 Human polypeptide
26	540.5	37.3	578	22	AA19502 Human polypeptide
27	529	36.5	415	21	AA19502 Human polypeptide
28	529	36.5	433	21	AA19502 Human polypeptide
29	529	36.5	441	21	AA19502 Human polypeptide
30	529	36.5	443	21	AA19502 Human polypeptide
31	529	36.5	461	21	AA19502 Human polypeptide
32	529	36.5	469	21	AA19502 Human polypeptide
33	526.5	36.3	540	21	AA19502 Human polypeptide
34	518	35.7	477	21	AA19502 Human polypeptide
35	518	35.7	500	21	AA19502 Human polypeptide
36	517.5	35.7	466	21	AA19502 Human polypeptide
37	507.5	35.0	476	21	AA19502 Human polypeptide
38	507.5	35.0	498	21	AA19502 Human polypeptide
39	507.5	35.0	499	21	AA19502 Human polypeptide
40	507.5	35.0	521	21	AA19502 Human polypeptide
41	507	35.0	465	21	AA19502 Human polypeptide
42	507	35.0	487	21	AA19502 Human polypeptide
43	490.5	33.8	274	21	AA19502 Human polypeptide
44	490.5	33.8	292	21	AA19502 Human polypeptide
45	490.5	33.8	300	21	AA19502 Human polypeptide

ALIGNMENTS

RESULT 1	AA19502	AA19502 standard; Protein; 277 AA.
ID	AA19502	
XX	AA19502	
AC	AA19502	
XX	AA19502	
DT	22-OCT-2001	(first entry)
XX	22-OCT-2001	
DE	Human polypeptide SEQ ID NO 2647.	
XX	Human; noctropic; immunosuppressant; cytostatic; gene therapy; cancer;	
KW	peripheral nervous system; neuropathy; central nervous system; CNS;	
KW	Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;	
KW	amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemocarcin;	
KW	chemokinetic; thrombolytic; drug screening; arthritis; inflammation;	
KW	leukaemia.	
XX		
OS	Homo sapiens.	
XX		
PN	WO200153312-A1.	
XX		
PD	26-UTL-2001.	
XX		
XX	26-DEC-2000; 2000WO-US4263.	
PF	21-JAN-2000; 2000US-0488725.	
XX	25-APR-2000; 2000US-0552317.	
PR	09-UTL-2000; 2000US-0598042.	
PR	19-UTL-2000; 2000US-0620312.	
PR	03-AUG-2000; 2000US-0653450.	
PR	14-SEP-2000; 2000US-0662191.	
PR	19-OCT-2000; 2000US-0693036.	

PR 29-NOV-2000; 2000US-0727344.
 XX (HYSE-) HYSEQ INC.
 XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
 XX WPI; 2001-442253/47.
 DR N-PSDB; AA158658.
 XX Novel nucleic acids and polypeptides, useful for treating disorders
 PT such as central nervous system injuries -
 PT Example 4; SEQ ID NO 2647; 10078pp; English.
 XX The invention relates to human nucleic acids (AA157798-AA161369) and
 CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
 CC immunosuppressant and cyostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemias and
 CC C.N.S disorders.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification.
 XX Sequence 277 AA;
 SQ Query Match 100.0%; Score 1450; DB 22; Length 277;
 Best Local Similarity 100.0%; Pred. No. 3,2e-132;
 Matches 277; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MASAGGEDCSPAPADRPORPFLIGVSGGTASGKSTVCCKIMELLGONEVEQORRKYV 60
 DB 1 MASAGGEDCSPAPADRPORPFLIGVSGGTASGKSTVCCKIMELLGONEVEQORRKYV 60
 QY 61 ILSODRFYKVLTAEOAKALKGOYNFDPDAFNDMLMHTLKNIVGKTVVEPTVDFVTH 120
 DB 61 ILSODRFYKVLTAEOAKALKGOYNFDPDAFNDMLMHTLKNIVGKTVVEPTVDFVTH 120
 QY 121 SRLPETTVVYPADVLFEGILVFYSQEI RDMFHLRLFVDTSDVRLSRVLRDVRGRDL 180
 DB 121 SRLPETTVVYPADVLFEGILVFYSQEI RDMFHLRLFVDTSDVRLSRVLRDVRGRDL 180
 QY 181 EQILTYTTTVPKPAFEFCFPTKKYADVILIPRGVDNMAINALIVQHIDILNGDICKMHR 240
 DB 181 EQILTYTTTVPKPAFEFCFPTKKYADVILIPRGVDNMAINALIVQHIDILNGDICKMHR 240
 QY 241 GGSNGRSYKRTFSEPGDHPGLTSGKSHLESSSRPH 277
 DB 241 GGSNGRSYKRTFSEPGDHPGLTSGKSHLESSSRPH 277
 RESULT 2
 AAB93941
 ID AAB93941 standard; Protein; 277 AA.
 XX AAB93941;
 AC 26-JUN-2001 (first entry)
 DT 26-JUN-2001 (first entry)
 XX Human protein sequence SEQ ID NO:13952.
 DE Human, primer; detection; diagnosis; antisense therapy; gene therapy.
 XX Homo sapiens.

XX EPI074617-A2.
 PN 07-FEB-2001.
 XX 28-JUL-2000; 2000EP-0116126.
 XX 29-JUL-1999; 93JP-0248036.
 PR 27-AUG-1999; 99JP-0300253.
 PR 11-JAN-2000; 2000JP-0118776.
 PR 02-MAY-2000; 2000JP-0183767.
 PR 09-JUN-2000; 2000JP-0241899.
 XX (HELI-) HELIX RES INST.
 PA Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 PI WPI; 2001-318749/34.
 DR The present invention describes primer sets for synthesising 5602
 XX full-length cDNAs defined in the specification, and for the detection
 PT and/or diagnosis of the abnormality of the proteins encoded by the
 PT full-length cDNAs -
 XX Claim 8; SEQ ID 13952; 2537pp + CD ROM; English.
 PS The present invention describes primer sets for synthesising 5602
 CC full-length cDNAs defined in the specification. Where a primer set
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
 CC to the complementary strand of a polynucleotide which comprises one of
 CC the 5602 nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in
 CC the specification. The primer sets can be used in antisense therapy and
 CC in gene therapy. The primers are useful for synthesising polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH1633 to AAH18742 represent human cDNA sequences; AAH95446 to
 CC AAH95893 represent human amino acid sequences; and AAH13629 to AAH1632
 CC represent oligonucleotides, all of which are used in the exemplification
 CC of the present invention.
 XX Sequence 277 AA;
 SQ Query Match 100.0%; Score 1450; DB 22; Length 277;
 Best Local Similarity 100.0%; Pred. No. 3,2e-132;
 Matches 277; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MASAGGEDCSPAPADRPORPFLIGVSGGTASGKSTVCCKIMELLGONEVEQORRKYV 60
 DB 1 MASAGGEDCSPAPADRPORPFLIGVSGGTASGKSTVCCKIMELLGONEVEQORRKYV 60
 QY 61 ILSODRFYKVLTAEOAKALKGOYNFDPDAFNDMLMHTLKNIVGKTVVEPTVDFVTH 120
 DB 61 ILSODRFYKVLTAEOAKALKGOYNFDPDAFNDMLMHTLKNIVGKTVVEPTVDFVTH 120
 QY 121 SRLPETTVVYPADVLFEGILVFYSQEI RDMFHLRLFVDTSDVRLSRVLRDVRGRDL 180
 DB 121 SRLPETTVVYPADVLFEGILVFYSQEI RDMFHLRLFVDTSDVRLSRVLRDVRGRDL 180
 QY 181 EQILTYTTTVPKPAFEFCFPTKKYADVILIPRGVDNMAINALIVQHIDILNGDICKMHR 240
 DB 181 EQILTYTTTVPKPAFEFCFPTKKYADVILIPRGVDNMAINALIVQHIDILNGDICKMHR 240
 QY 241 GGSNGRSYKRTFSEPGDHPGLTSGKSHLESSSRPH 277

DT 22-OCT-2001 (first entry)
XX
DE Human polypeptide SEQ ID NO 6219.
XX
XX Human, noctropic; immunosuppressant; cytostatic; gene therapy; cancer;
KM peripheral nervous system; neuropathy; central nervous system; CNS;
KM Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KM amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KM chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KM leukaemia.
XX
OS Homo sapiens.
XX
PN WO200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US34263.
XX
PR 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0683036.
PR 29-NOV-2000; 2000US-0727344.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Aunndi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou F, Goodrich R, Drmanac RT;
XX
XX WPI; 2001-442253/47.
DR N-PSDB; AAI60444.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
XX
XX Example 2; SEQ ID NO 6219; 10078bp; English.
XX
XX The invention relates to human nucleic acids (AA157798-AA161369) and
CC the encoded polypeptides (AAM38642-AAM4213) with noctropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX
XX
SQ Sequence 296 AA;
Query Match 100.0%; Score 1450; DB 22; Length 296;
Best Local Similarity 100.0%; Pred. No. 3.5e-132;
Matches 277; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MASAGGEDCESPAEADRPORPFLIGVSGTASGKSTVCEKIMELLGONEVEORQRKV 60
DB 20 MASAGGEDCESPAEADRPORPFLIGVSGTASGKSTVCEKIMELLGONEVEORQRKV 79
QY 61 ILSODRFYKVLTAEOAKALKGOYNFDHPDAFNDMLMRTLKNIIVEGKTVVPTDYFVTH 120
DB 80 ILSODRFYKVLTAEOAKALKGOYNFDHPDAFNDMLMRTLKNIIVEGKTVVPTDYFVTH 139

QY 121 SRLPETTVYPADVLFEGILVFSOEIRDMFRLRLFVTDSDVRLSRVLRDVRGRD 180
DB 140 SRLPETTVYPADVLFEGILVFSOEIRDMFRLRLFVTDSDVRLSRVLRDVRGRD 199
QY 181 EQLITQYTFVFKPAFEFCPLTKKYADVILPRGVDMVAIINLIVQHIOILNGDICKMR 240
DB 200 EQLITQYTFVFKPAFEFCPLTKKYADVILPRGVDMVAIINLIVQHIOILNGDICKMR 259
QY 241 GGSNGRSYKRTFSEPGDHPGMLTSGKSHLESSSRH 277
DB 260 GGSNGRSYKRTFSEPGDHPGMLTSGKSHLESSSRH 296
RESULT 5
ID AAG64506 standard; Protein; 276 AA.
XX
XX AAG64506;
AC
XX 02-OCT-2001 (first entry)
DT
XX Human uridine kinase.
DE
XX Human, uridine kinase; UK.
XX
OS Homo sapiens.
XX
PN CN1287172-A.
XX
PD 14-MAR-2001.
XX
PF 07-SEP-1999; 99CN-0118818.
PR 07-SEP-1999; 99CN-0118818.
XX
XX (UYFU-) UNIV FUDAN.
PA
PI Yu L, Zhao Y, Zhang H;
PI
XX WPI; 2001-409529/44.
DR N-PSDB; AAH75355.
XX
PT Human uridine kinase and its coding sequence, preparation and
PT application -
XX
XX Claim 2; Page 15-16 (Disclosure); 20pp; Chinese.
XX
XX The invention relates to human uridine kinase (UK).
XX
SQ Sequence 276 AA;
Query Match 94.7%; Score 1372.5; DB 22; Length 276;
Best Local Similarity 96.0%; Pred. No. 1e-124;
Matches 266; Conservative 3; Mismatches 7; Indels 1; Gaps 1;
QY 1 MASAGGEDCESPAEADRPORPFLIGVSGTASGKSTVCEKIMELLGONEVEORQRKV 60
DB 1 MASAGGEDCESPAEADRPORPFLIGVSGTASGKSTVCEKIMELLGONEVEORQRKV 59
QY 61 ILSODRFYKVLTAEOAKALKGOYNFDHPDAFNDMLMRTLKNIIVEGKTVVPTDYFVTH 120
DB 60 ILSODRFYKVLTAEOAKALKGOYNFDHPDAFNDMLMRTLKNIIVEGKTVVPTDYFVTH 119
QY 121 SRLPETTVYPADVLFEGILVFSOEIRDMFRLRLFVTDSDVRLSRVLRDVRGRD 180
DB 120 SRLPETTVYPADVLFEGILVFSOEIRDMFRLRLFVTDSDVRLSRVLRDVRGRD 179
QY 181 EQLITQYTFVFKPAFEFCPLTKKYADVILPRGVDMVAIINLIVQHIOILNGDICKMR 240
DB 180 EQLITQYTFVFKPAFEFCPLTKKYADVILPRGVDMVAIINLIVQHIOILNGDICKMR 239
QY 241 GGSNGRSYKRTFSEPGDHPGMLTSGKSHLESSSRH 277

CC immunological disorders. This sequence represents the protein relating to
 CC a human uridine kinase (UDK) of the invention.
 CC NOTE: The present sequence is stated as being the same as that shown as
 CC SEQ ID NO. 2 in the sequence listing of the specification. However, the
 CC sequences differ.

XX Sequence 277 AA;

Query Match 91.4%; Score 1325; DB 22; Length 277;
 Best Local Similarity 91.7%; Pred. No. 4,3e-126;
 Matches 254; Conservative 11; Mismatches 12; Indels 0; Gaps 0;

QY 1 MASAGDECSPEADRPHPORPFLIGVSGGTASGKSTVCEKIMELLGONEVEQORQKV 60
 DB 1 MASAGGSESADPEADRPORPFLIGVSGGTASGKSTVCEKIMELLGONEVEQORQKV 60
 QY 61 ILSODRFKVLTAEOKAKALKGQYFNDHPDAFNDMLMHTLKNIVEGTVVEPTYPVTH 120
 DB 61 ILSODCFKVLTAEOKAKALKGQYFNDHPDAFNDMLMHTLKNIVEGTVVEPTYPVTH 120
 QY 121 SRLPETTVVPADVLPFGILVFSQEI RDMFHLRLFPVDTSDVRLSRVLRDVRGRDL 180
 DB 121 SRLPETTVVPADVLPFGILVFTIQEIRDMFHLRLFPVDTSDVRLSRVLRDVRGRDL 180
 QY 181 EQILTOYTTFVKPAFEBCLEPTKKYADYIIPRGYDNNVAIINLYQHIQDILNGDLCKMR 240
 DB 181 EQILTOYTAFAVKPAFEBCLEPTKKYADYIIPRGYDNNVAIINLYQHIQDILNGDLCKMR 240
 QY 241 GGSNGRSKRTFSEPRGDHPGMLTSGKSHLESSSRPH 277
 DB 241 GGPNGRNHKTFFPEPGDHGVLATGKSHLESSSRPH 277

RESULT 8

ID ABB89353 standard; Protein; 190 AA.

XX ABB89353;

DT 24-MAY-2002 (first entry)

DE Human polypeptide SEQ ID NO 1729.

XX Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
 KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor;
 KW vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
 KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; human; secreted protein.

XX Homo sapiens.

OS WO200190304-A2.

XX 29-NOV-2001.

PF 18-MAY-2001; 2001WO-US16450.

XX 19-MAY-2000; 2000US-20551SP.

PA (HUMA-) HUMAN GENOME SCI INC.

PI Birse CE, Rosen CA;

DR WPI: 2002-122016/16.

DR N-PSDB; ABL89762.

XX Novel 1405 isolated polypeptides, useful for diagnosis, treatment and
 PT prevention of neural, immune system, muscular, reproductive,
 PT gastrointestinal, pulmonary, cardiovascular, renal and proliferative
 disorders -

PS Claim 11; SEQ ID NO 1729; 2081IP + Sequence Listing; English.

XX

CC The invention relates to novel genes (ABL89449-ABL90853) and proteins
 CC (ABB89040-ABB90444) useful for preventing, treating or ameliorating
 CC medical conditions e.g. by protein or gene therapy. The genes are
 CC isolated from a range of human tissues disclosed in the specification.
 CC The nucleic acids, proteins, antibodies and (ant)agonists are useful
 CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
 CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
 CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
 CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
 CC hemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
 CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
 CC colitis; (c) cardiovascular disorders such as myocardial ischaemia;
 CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
 CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
 CC and parasitic infections.

CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at fcp.wipo.int/pub/published_pct_sequences.

XX Sequence 190 AA;

Query Match 64.9%; Score 941; DB 23; Length 190;
 Best Local Similarity 98.9%; Pred. No. 4,6e-83;
 Matches 183; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MASAGDECSPEADRPHPORPFLIGVSGGTASGKSTVCEKIMELLGONEVEQORQKV 60
 DB 1 MASAGDECSPEADRPHPORPFLIGVSGGTASGKSTVCEKIMELLGONEVEQORQKV 60

QY 61 ILSODRFKVLTAEOKAKALKGQYFNDHPDAFNDMLMHTLKNIVEGTVVEPTYPVTH 120
 DB 61 ILSODRFKVLTAEOKAKALKGQYFNDHPDAFNDMLMHTLKNIVEGTVVEPTYPVTH 120

QY 121 SRLPETTVVPADVLPFGILVFSQEI RDMFHLRLFPVDTSDVRLSRVLRDVRGRDL 180
 DB 121 SRLPETTVVPADVLPFGILVFTIQEIRDMFHLRLFPVDTSDVRLSRVLRDVRGRDL 180

QY 181 EQILTOYTTFVKPAFEBCLEPTKKYADYIIPRGYDNNVAIINLYQHIQDILNGDLCKMR 240
 DB 181 EQILTOYTTFVKPAFEBCLEPTKKYADYIIPRGYDNNVAIINLYQHIQDILNGDLCKMR 240

QY 241 GGSNGRSKRTFSEPRGDHPGMLTSGKSHLESSSRPH 277
 DB 241 GGPNGRNHKTFFPEPGDHGVLATGKSHLESSSRPH 277

QY 181 EQILT 185

DB 181 EQILT 185

QY 181 EQILT 185

DB 181 EQILT 185

QY 181 EQILT 185

DB 181 EQILT 185

QY 181 EQILT 185

DB 181 EQILT 185

QY 181 EQILT 185

DB 181 EQILT 185

QY 181 EQILT 185

DB 181 EQILT 185

QY 181 EQILT 185

DB 181 EQILT 185

QY 181 EQILT 185

DB 181 EQILT 185

QY 181 EQILT 185

DB 181 EQILT 185

QY 181 EQILT 185

DB 181 EQILT 185

QY 181 EQILT 185

DB 181 EQILT 185

PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
XX
PA (HYSE-) HYSEO INC.
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX WPI; 2001-442253/47.
DR N-PSDB; AA157850.
XX
XX Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
XX
XX Example 3; SEQ ID NO 1839; 10078bp; English.
PS
XX The invention relates to human nucleic acids (AA157798-AA161369) and
CC the encoded polypeptides (AA38642-AA42213) with nootropic,
CC immunosuppressant and cyostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX
SQ Sequence 261 AA;
Query Match 64.4%; Score 934; DB 22; Length 261;
Best Local Similarity 69.5%; Pred. No. 3,4e-82;
Matches 189; Conservative 27; Mismatches 40; Indels 16; Gaps 4;
QY 8 DCEPPAEADRPH-QRPFLIGVSGGTASGKSTVCEKIMELLGONEVEBQRKRVILISQDR 66
DB 4 DSEQTQNHQONGGEPFLIGVSGGTASGKSSVCARKIVQLLGONEVDYRKQVILISQDS 63
QY 67 FYKVLTRFQAKALKGQYNDHPDAPFNDMLMHTLKNIVGKYEVVTYTFVTHSRLEPT 126
DB 64 FYEVLSEQAKALKGQFNDHPDAPFNEMLLTKLKEITGKTVQIFVYFVSHSRKEET 123
QY 127 TVVYPADVVFEGILVYVYSGOIRDMFRLRFVPTDSVRLSRVRLRPDV-RRGRDLRQILT 185
DB 124 VVYVPADVVFEGILAFYVYSGOIRDLRFQMKLFVDIDATRLSRVRLDISRGDLRQILS 183
QY 186 QYTFVVPAREEFCLEPTKAYADVIIPRGVDMVAIINLVQIHODILNGDICKMHRGSSNG 245
DB 184 QYITFVVPAREEFCLEPTKAYADVIIPRGADNLVAIINLVQIHQIDILNG-----G 232
QY 246 KSTYKTFSEPGDHPGMLTSGKSHLSSSRPH 277
DB 233 PSKRQTF---NGCLNGYTPSRKRQASESSSRPH 261
RESULT 10
AAB73494
ID AAB73494 standard; Protein; 261 AA.
XX
XX AAB73494;
AC
XX
XX 31-JUL-2001 (first entry)
DT
XX Human transferase HTFS-1, SEQ ID NO:1.
DE
XX Human transferase HTFS-1, SEQ ID NO:1.
KW Human transferase; HTFS; agonist; antagonist; cellular signalling;
KW proliferation; cell proliferative disorder; immune disorder;

KW atherosclerosis; hepatitis; psoriasis; cancer; tumour;
KW inflammation; AIDS; Addison's disease; allergy; asthma; anaemia;
KW cirrhosis; Crohn's disease; atopic dermatitis; diabetes mellitus;
KW multiple sclerosis; rheumatoid arthritis; pancreatitis;
KW systemic lupus erythematosus; thrombocytopenia; ulcerative colitis;
KW haemodialysis; extracorporeal circulation; trauma; transgenic animal;
KW gene therapy; drug screening.
XX
OS Homo sapiens.
XX
XX WO200132888-A2.
PN
XX 10-MAY-2001.
PD
XX 02-NOV-2000; 2000WO-US30485.
PF
XX 04-NOV-1999; 99US-0163595.
PR
XX (INCYTE GENOMICS INC.
PA
XX Tang YT, Yue H, Hillman UL, Lal P, Bandman O, Patterson C;
PI Shih LL, Azimtai Y, Lu DAM, Baughn WR;
XX WPI; 2001-328796/34.
DR N-PSDB; AA423801.
DR
XX Human transferase polypeptides and polynucleotides useful for
PT diagnosis, prevention and treatment of cell proliferative and immune
PT system disorders and for identifying agonists and antagonists -
XX
PS Claim 1; Page 103-104; 157pp; English.
XX
XX Sequences AAB73494-AAB73335 represent novel human transferase proteins
CC HTFS-1 to HTFS-42, and sequences AA423801-AA423842 represent cDNAs
CC encoding them. The proteins play important roles in the regulation of
CC cellular signalling and proliferation. The HTFS proteins are useful for
CC screening compounds for their effectiveness as agonists or antagonists of
CC transferase activity, or for compounds that specifically bind to an HTFS
CC protein or which modulates the activity of an HTFS protein.
CC Pharmaceutical compositions comprising an HTFS protein, HTFS
CC agonist or antagonist, or genetic construct encoding an HTFS
CC protein are useful for treating a disease or condition associated
CC with decreased or increased expression of functional HTFS. Disorders
CC which may be treated using such compositions include cell proliferative
CC disorders and immune disorders. For example, diseases which may be
CC treated include atherosclerosis, hepatitis, psoriasis, cancers (including
CC breast, bladder, bone marrow, brain and uterus cancer), inflammation,
CC AIDS, Addison's disease, allergies, asthma, anaemia, cirrhosis, Crohn's
CC disease, atopic dermatitis, diabetes mellitus, multiple sclerosis,
CC rheumatoid arthritis, pancreatitis, systemic lupus erythematosus,
CC thrombocytopenia, and ulcerative colitis. They may also be used to treat
CC complications of cancer, haemodialysis, extracorporeal circulation,
CC trauma and haematopoietic cancer, including lymphoma, leukaemia and
CC myeloma. Polynucleotides encoding HTFS proteins are useful for creating
CC transgenic animals to model human diseases, for diagnostic purposes and
CC to generate hybridisation probes useful in mapping the naturally
CC occurring genomic sequences. HTFS, and its catalytic or immunogenic
CC fragments are useful for screening libraries of compounds in a variety of
CC drug screening techniques. Antibodies which specifically bind HTFS may be
CC used for the diagnosis of disorders associated with the expression of
CC HTFS, or in assays to monitor patients being treated with HTFS or
CC agonists, antagonists or inhibitors of HTFS. The present sequence
CC represents an HTFS protein of the invention.
XX
SQ Sequence 261 AA;
Query Match 64.4%; Score 934; DB 22; Length 261;
Best Local Similarity 69.5%; Pred. No. 3,4e-82;
Matches 189; Conservative 27; Mismatches 40; Indels 16; Gaps 4;
QY 8 DCEPPAEADRPH-QRPFLIGVSGGTASGKSTVCEKIMELLGONEVEBQRKRVILISQDR 66
DB 4 DSEQTQNHQONGGEPFLIGVSGGTASGKSSVCARKIVQLLGONEVDYRKQVILISQDS 63

XX 08-MAR-2000; 2000WO-US05988.
PF 12-MAR-1999; 99US-0124270.
XX (HUMA-) HUMAN GENOME SCI INC.
PA (ROSE/) ROSEN C A.
XX
PI Rosen CA, Ruben SM,
XX WPI; 2000-587513/55.
DR N-PSDB; AAF15785.
XX
PT Prostate cancer associated gene sequences, referred to as prostate
PT cancer antigens, useful for treatment, prevention, and diagnosis of
PT disorders such as prostate cancer -
XX
XX Claim 11; Page 1566-1567; 2338pp; English.
XX
CC AAF15566 to AAF16505 encode the human prostate cancer associated
CC proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.
CC The prostate cancer antigens can have neuroprotective, cytostatic,
CC cardioprotective, immunomodulatory, muscular, vulnary, gastrointestinal,
CC nephrotoxic, anti-infective, gynaecological and antibacterial activities,
CC and can be used in gene therapy. The prostate cancer antigen
CC polynucleotides may be used for detection of prostate cancer, chromosome
CC identification, as chromosome markers, and for numerous other diagnostic
CC or research purposes. The prostate cancer antigens may be used to treat
CC disorders such as neural, immune, muscular, reproductive,
CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
CC disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to
CC AAB57303 represent sequences used in the exemplification of the present
CC invention.
XX
SQ Sequence 337 AA;

Query Match 61.8%; Score 896; DB 21; Length 337;
Best Local Similarity 62.3%; Pred. No. 2,3e-78;
Matches 188; Conservative 28; Mismatches 40; Indels 46; Gaps 6;

QY 2 ASAGGE-----DCESPAPADR--PHQRPFLIGVSGTASGK 36
DB 56 AGAGGERASVRTSGRRGANHGRCQADPAEPAPARRALPYRR-----HGTASGK 109
QY 37 STYCEKIMELLIGQNEVQRQKRVILSODRFYKULTBQKAKLKQYNDPDPAPFNDL 96
DB 110 SSVCATKIVOLLGQNEVYRQKVILSQDSFYRLTSEOKAKALKQFNFPHPAPFNMEX 169
QY 97 MHRTLKIVGKTEVPTVDFVTHSRLEPTTVVPADVLPFEGILVFSOEIRMFHLRL 156
DB 170 ILTKLKEITBEKIVQIVDFVSHSRKEBTYVTPADVLPFEGILAFYSQEVRLDFQMKL 229
QY 157 FVPTDSDVRLSRVLRDV-RRGRDLEOILTQYTFVAPAEFECLPTKXADAVILPRGVD 215
DB 230 FVPTDADTRLSRVLDRDISERGRDLEQILSYITFVAPAEFECLPTKXADAVILPRGAD 289
QY 216 NMVAINILVQHIOILLGDLCKMHRGSGNSRYKRTSEPDHMGMLTSGKSHLESSR 275
DB 290 NLVAINILVQHIOILLG-----GPSKRQT---NGCLNGYTPSRKRFQASESSSR 335
QY 276 PH 277
DB 336 PH 337

RESULT 13
ID AAM40480
AC AAM40480;
DT 22-OCT-2001 (first entry)

DE Human polypeptide SEQ ID NO 5411.

XX Human, nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KM peripheral nervous system; neuropathy; central nervous system; CNS;
KM Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KM amyotrophic lateral sclerosis; Shy-Drager Syndrome; Chemotactic;
KM chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KM leukaemia.
XX
OS Homo sapiens.
XX
PN WO20015312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US34263.
XX
XX 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
DR WPI; 2001-442253/47.
DR N-PSDB; AAI59636.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
XX
PS Example 2; SEQ ID NO 5411; 10078pp; English.
XX
CC The invention relates to human nucleic acids (AA157798-AA161369) and
CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC immunosuppressant and cyostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemia and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX
SQ Sequence 335 AA;

Query Match 61.3%; Score 889; DB 22; Length 335;
Best Local Similarity 63.8%; Pred. No. 1,1e-77;
Matches 190; Conservative 26; Mismatches 46; Indels 36; Gaps 7;

QY 2 ASAGGE-----DCESPAPADRPH-ORPFLIGVSGTASGKSTVCEK 42
DB 52 AGAGGERASVRTSGRRGRTMAGDSEQTLQNHQPNCGEPFLIGVSGTASGKSSVCAC 111
QY 43 IMELLGQNEVQRQKRVILSODRFYKULTBQKAKLKQYNDPDPAPFNDLMHRTLK 102
DB 112 IVOLLGQNEVYRQKVILSQDSFYRLTSEOKAKALKQFNFPHPAPFNDLILTKLK 171
QY 103 NIVEGKTEVPTVDFVTHSRLEPTTVVPADVLPFEGILVFSQE-IRDMFHLRLFVPTD 161

```

Dh 172 ETEGKTQVQIPVYDFVSHSRKETVTVPADVLFEGILAFYSQERIRDLFQMKLFVDTD 231
Qy 162 SDVLSRRVLRDV-RRGDLEQILLTOYT-TVKPAFEFCF.PTKKYADVILPRGVDMVA 219
Dh 232 ADTLSRRVLRDVLSERGRDLEQILLSSSTLRFPKPAFEFCF.PPKKYADVILPRGADNRVP 291
Qy 220 INLIYVQHIDILNDICKMHGSGNSGRSYKRTFSEPDGHPGMLTSGKSHLESSSRPH 277
Dh 292 INLIYVQHIDILNG-----GPNRQT---NGCLNGYTRSRKRGASSESSSRPH 335

RESULT 14
ABG27216
ID ABG27216 standard; Protein; 260 AA.
XX
AC ABG27216;
XX
DT 26-FEB-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 13713.
XX
KM Drosophila; developmental biology; cell signalling; insecticide;
XX pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
XX
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE ) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-6556860/75.
XX
DR N-PSDB; ABL06410.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions -
XX
PS Disclosure; SEQ ID NO 13713; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
XX sequences (AB101840-AB16175) and the encoded proteins
XX (ABB57737-ABB72072).
XX
CC The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 260 AA;
XX
Query Match 50.9%; Score 738; DB 22; Length 260;
Best Local Similarity 67.8%; Pred. No. 3.5e-63;
Matches 143; Conservative 32; Mismatches 34; Indels 2; Gaps 2;
Qy 23 PFLIGVSGTASGKSTVCEKIMELIGONEVEORQRKVILISODRFYVLTAQKAKALG 82
Dh 27 PFLIGVAGTASGKSTVCKKIMEQLGAEQMDHTRQVVSISQDSFYRELTPAEKAKXKG 86
Qy 83 QYNFDDHDAFNDMLMARTLKNIYEGKTVVEVTFYVTHS-RLPFTTVYVPADVILFEGL 141
Dh 87 LFNFDHDAFNEELMYSTLQNLKGHKVEIPSYDRTNSLDFENVLVIYPADVILFEGL 146

```

```

Qy 142 VFYSOERDMFHLRLFYVDTSDVRLSRVLRDV-RRGDLEQILLTOYTTPVKPAFEFCF 200
Dh 147 VFYFPKRLRFHMKLFYVDTSDTRLRARRVFRDINEGRDLDAVLTQYMTYVKPAFEFCF 206
Qy 201 PTKKYADVILPRGVDMVAINLIYVQHIDIL 231
Dh 207 PTKKYADVILPRGADNTVAIDLIVHHIGEIL 237

RESULT 15
ABG27216
ID ABG27216 standard; Protein; 326 AA.
XX
AC ABG27216;
XX
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #27207.
XX
KM Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
XX
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
XX
DR N-PSDB; AAS91403.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX PT responsible for genetic disorders or other traits and to assess
XX biodiversity -
XX
PS Claim 20; SEQ ID NO 57575; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
XX polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX and gene mapping, and in recombinant production of (II). The
XX CC polynucleotides are also used in diagnostics as expressed sequence tags
XX for identifying expressed genes. (I) is useful in gene therapy techniques
XX CC to restore normal activity of (II) or to treat disease states involving
XX (II). (II) is useful for generating antibodies against it, detecting or
XX quantitating a polypeptide in tissue, as molecular weight markers and as
XX a food supplement. (II) and its binding partners are useful in medical
XX CC imaging of sites expressing (II). (I) and (II) are useful for treating
XX disorders involving aberrant protein expression or biological activity.
XX CC The polypeptide and polynucleotide sequences have applications in
XX CC diagnostics, forensics, gene mapping, identification of mutations
XX and to produce other types of data and products dependent on DNA and
XX CC amino acid sequences. ABG00010-ABG30377 represent novel human
XX CC Note: The sequence data for this patent did not appear in the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 326 AA;
XX
Query Match 40.6%; Score 589; DB 22; Length 326;

```


This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: November 25, 2003, 10:07:48 ; Search time 2539 Seconds

(without alignments)
2651.571 Million cell updates/sec

29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Title: US-09-896-522-2
Perfect score: 277
Sequence: 1 MASAGEDESPAPRADRP.....HPGMLTSGKSHLESSRPH 277

Scoring table:
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 22781392 seqs, 12152238056 residues

Word size: 15

Total number of hits satisfying chosen parameters: 509

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n.model -DB=xlh
-Q=/cgn2_1/USPTO.spool/US09896522/rnuc 21112003_184459_5876/app_query.fasta.1.455
-DB=EST -QPM=faetap -SUFFIX=p2n_ol1.rst -MINMATCH=0.1 -LOOPL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=quality -THR MIN=15 -ALIGN=15 -MODE=LOCAL -OUTFMT=pro
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09896522.@CGN_1.1.2810.@runac 21112003_184459_5876 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : EST.*

1: em_estba.*
2: em_esthum.*
3: em_estin.*
4: em_estnu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_hic.*
9: gb_est1.*
10: gb_est2.*
11: gb_hic.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: em_gss_hum.*
18: em_gss_inv.*
19: em_gss_pln.*
20: em_gss_vtl.*
21: em_gss_fun.*
22: em_gss_mam.*
23: em_gss_mus.*
24: em_gss_pro.*
25: em_gss_rtd.*
26: em_gss_png.*
27: em_gss_vtl.*
28: gb_gss1.*

Result No.	Score	Query Match	Length	DB ID	Description
1	236	85.2	1044	12	BM917506
2	231	83.4	1201	12	BX394295
3	230	83.0	1118	12	BM467984
4	228	82.3	898	14	CD514811
5	225	81.2	725	12	B1256928
6	216	78.0	916	12	BG826894
7	214	77.3	1036	13	BQ072501
8	213	76.9	730	12	BG770518
9	213	76.9	869	10	BG491384
10	213	76.9	956	13	BQ719741
11	208	75.1	989	10	BG390519
12	205	74.0	827	12	B1261258
13	203	73.3	860	10	BG491358
14	200	72.2	1039	12	BM803129
15	197	71.1	848	12	B1258532
16	196	70.8	810	12	B1753007
17	194	70.0	1055	12	BM545603
18	193	69.7	661	12	B1830878
19	193	69.7	910	13	BQ437115
20	192	69.3	922	13	BQ935919
21	191	69.0	753	9	AU122008
22	180	68.6	963	14	CD359372
23	185	66.8	959	12	B1770572
24	185	66.8	1095	12	BM802939
25	182	65.7	778	14	CD101693
26	177	63.9	554	12	BM822058
27	174	62.8	584	10	AW961647
28	169	61.0	799	10	BG723578
29	169	61.0	1114	13	BX343101
30	168	60.6	998	13	BX400889
31	163	58.8	514	12	BM827668
32	163	58.8	515	12	BM835278
33	160	57.8	585	12	CB156232
34	160	57.8	923	10	BB891861
35	157	56.7	945	13	BQ677293
36	156	56.3	883	13	BQ838101
37	156	56.3	948	13	BQ680950
38	156	56.3	973	13	BQ683188
39	153	55.2	797	12	B1116432
40	153	55.2	898	13	BQ680187
41	152	54.9	1097	10	BF347745
42	150	54.2	819	12	B1910015
43	146	52.7	494	9	AU129365
44	145	52.3	819	12	B1770639
45	141	50.9	425	9	A1640490

ALIGNMENTS

RESULT 1
LOCUS BM917506 1044 bp mRNA linear EST 12-MAR-2002
DEFINITION AGENCOURT 6606568 NIH_MGC_106 Homo sapiens CDNA clone IMAGE:5484033
ACCESSION BM917506
VERSION 5', mRNA sequence.
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 1044)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: Dr. Daniel McVicar, DBS/NCI
 CDNA Library Preparation: Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
 Plate: LNCM2012 row: k column: 10
 High quality sequence stop: 624.

FEATURES
 Source
 1. 1044
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5484033"
 /tissue_type="natural killer cells, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC 106"
 /note="Organ: blood; Vector: pOT87; Site 1: XhoI; Site 2:
 EcoRI; CDNA made by oligo-dT priming. Directionally cloned
 into EcoRI/XhoI sites using the following 5' adaptor:
 GGACGAG(G). Library constructed by Ling Hong in the
 Laboratory of Gerald M. Rubin (University of California,
 Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and
 Superscript II RT (Life Technologies). Note: this is a
 NIH MGC Library."

BASE COUNT 224 a 306 c 318 g 196 t

ORIGIN

Alignment Scores:
 Pred. No.: 5.77e-226 Length: 1044
 Score: 236.00 Matches: 236
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 85.20% Indels: 0
 DB: 12 Gaps: 0

US-09-896-522-2 (1-277) x BM917506 (1-1044)

QY 5 GlyGlyGlyUAAPCyGUserProAlaProGluAlaAspArgProHisGlnArgProPhe 24
 Db 1 GGAGGCGGAAGACTCGGAGAGCCCGCGCGGAGGCGGACCTCCGACACGCGCCCTTC 60
 QY 25 LeuIleGlyValSerGlyGlyThrAlaSerGlyLysSerThrValCysGluLysIleMet 44
 Db 61 CTGATTAGGGGTGAGCGCGGCACTGCCAGCGGAGAGTTCGACGTTGTGACAAATCATG 120
 QY 45 GluLeuLeuGlyGlnAsnGluValGluGlnArgGlnArgLysValValIleLeuSerGln 64
 Db 121 GAGTTGCTGGGACAGAACGAGTGGAAACAGCGGAGCGGAGAGTGCATCTGAGCGAG 180
 QY 65 AspArgPheTyrLysValLeuThrAlaGluGlnLysAlaLysValAlaLeuLysGlyGlnTyr 84
 Db 181 GACAGGTTCTCAAGAGTCTTACGCGGACGAGAGGCGGAGGCTTGAAAGACAGTAC 240
 QY 85 AsnPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThrLeuLysAsnIle 104
 Db 241 AATTTTGAACATCCAGATGCTTTGATTAATATTGATGACACAGATCTCGAAGAACATC 300
 QY 105 ValGluGlyLysThrValGluValProThrTyrAspPheValThrHisSerArgLeuPro 124
 Db 301 GTGAGAGGCAAAAGCGTGAAGTGCACCTATGATTTTGGACACACTCAAGGTTTACA 360
 QY 125 GluThrThrValValTyrProAlaAspValValLeuPheGluGlyIleLeuValPheTyr 144
 Db 361 GAGAGCCACGGTGTCTACCTCGCGGACGTGTTCTGTTTGAGGCGATCTTGTTCTTAC 420
 QY 145 SerGlnIuIleArgAspMetPheHisLeuArgLeuPheValAspThrAspSerAspVal 164

Db 421 AGCAGAGATCCGGAGACATTTTCCACTGCGGCTTTGCTGACACCGACTCCGACGTC 480
 QY 165 ArgLeuSerArgArgValLeuArgAspValArgArgGlyArgAspLeuGlnIleLeu 184
 Db 481 AGGCTGCTCGAAGATTCTCCGGAGCGTGCAGGAGGAGGAGGAGGAGGAGGAGGAGG 540
 QY 185 ThrGlnTyrThrThrPheValLysProAlaPheGluGluPheCysLeuProThrLysLys 204
 Db 541 ACCGAGTACACCACTTCGTAAGCCGCGCTTGAGAGAGTTGCTGCGCAAAAGAG 600
 QY 205 TyrAlaAspValIleIleProArgGlyValAspAspMetValAlaIleAsnLeuIleVal 224
 Db 601 TATGCCGATGTATCTATCCACAGAGAGTGCAGAAATGTGTTCCATCACTGATCGTG 660
 QY 225 GlnHisIleGlnAspIleLeuAsnGlyAspIleCysLysTyrHisArg 240
 Db 661 CAGCAGATCCAGAGACATTCGAATGTTGACATCTGCAATGACACCGA 708

RESULT 2
 BX394295 1201 bp mRNA linear EST 13-MAY-2003
 LOCUS BX394295 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
 DEFINITION CDNA clone CS0DC013YE13 5-PRIME, mRNA sequence.
 ACCESSION BX394295 GI:30624219
 VERSION BX394295.1
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 1201)
 Full-length cDNA libraries and normalization
 Unpublished
 Contact: Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: segret@genoscope.cns.fr, Web: www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 4968.r For
 more information about this cluster, see
<http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0DC013AC07QPL&cluster=4968.r>
 Contact : Feng Liang Email: liang@lifetech.com URL: <http://fulllength.invitrogen.com/> Invitrogen Corporation 1600
 Faraday Avenue Genoscope sequence ID: CS0DC013AC07QPL.

FEATURES
 source
 1. 1201
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DC013YE13"
 /tissue_type="NEUROBLASTOMA COT 25-NORMALIZED"
 /clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
 /note="First strand cDNA was primed with a NotI-oligo(dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and cloned into the Not I and EcoR V
 sites of the pCMVSPORT 6 vector. Library was normalized."

BASE COUNT 288 a 284 c 363 g 226 t 40 others

ORIGIN

Alignment Scores:
 Pred. No.: 6.94e-221 Length: 1201
 Score: 231.00 Matches: 245
 Percent Similarity: 99.59% Conservative: 0
 Best Local Similarity: 99.59% Mismatches: 1
 Query Match: 83.39% Indels: 1
 DB: 13 Gaps: 0

US-09-896-522-2 (1-277) x BX394295 (1-1201)

QY 1 MetAlaSerAlaGlyGlyUAAPCyGUserProAlaProGluAlaAspArgProHis 20

```

Db      88 ATGCGCTTCGGCGGAGCGAAGACTGCGAGAGCCCGCGCGGA-GCCGACCGTCCGCAC 146
Qy      21 GlnArgProPheLeuIleGlyValSerGlyGlyThrAlaSerGlyLysSerThrValCys 40
Db      147 CACCGCGCCCTTCCTGATAGGGGTGAGCGCGGCACCTCCAGCGGGAAGTGCACCGTGTGT 206
Qy      41 GlnLysIleMetGlnLeuLeuGlnGlnAlaGlnGlnValGlnGlnArgGlnArgLysValVal 60
Db      207 GAGAGATCATGGAGTGTCTGGAGCAGAACGAGTGAACGAGCGCGCAGGAGAGTGTGT 266
Qy      61 IleLeuSerGlnAspArgPheTyrlsValLeuThrAlaGlnGlnLysAlaLysAlaLeu 80
Db      267 ATCTGAGCCAGAGCAGAGTCTCAACAGTCTTGAACGCGCAGAGCAGAGCGCAGCCTTG 326
Qy      81 LysGlyGlnTyrlsAspPheAspHisProAspAlaPheAspLysPheMetHisArgThr 100
Db      327 AAGGACAGTACATTTTGAACCATCCAGATGCCCTTTGATATGATTTGATGACAGGACT 386
Qy      101 LeuLysAsnIleValGlnGlyLysThrValGlnValProThrTyrlsAspPheValThrHis 120
Db      387 CTGAGAACATCTGTGAGGGGAAACGAGTGAAGTGCAGCTATGATTTGTGACACAC 446
Qy      121 SerArgLeuProGlnThrThrValValTyrlsProAlaAspValValLeuPheGlnGlyIle 140
Db      447 TCAAGGTACCAAGACCAAGGTGTCTACCTGCGGACGTGTTCTGTTGAGGGGATC 506
Qy      141 LeuValPheTyrlsSerGlnGlnIleArgAspMetPheHisLeuArgLeuPheValAspThr 160
Db      507 TTGGTGTCTTACAGCCAGAGATCCGGGACATGTTCCACCTGCGCCTTCTGTGACACC 566
Qy      161 AspSerAspValArgLeuSerArgArgValLeuArgAspValArgArgGlyArgAspLeu 180
Db      567 GACTCGACATCGAGCTGTCTGAGAGGTCTCCGGACGTGCGCAGGAGGAGGACTG 626
Qy      181 GlnGlnIleLeuThrGlnTyrlsThrThrPheValLysProAlaPheGlnGlnGlyLeu 200
Db      627 GACGAGTTCTGACGCGATCACCACTTGTGTAGAGCGGCTTTCAGAGAGTCTGTGCTG 666
Qy      201 ProThrLysLysTyrlsAlaAspValIleIleProArgGlyValLysAspMetValAlaIle 220
Db      687 CCGACAAAGAAATGCGGATGATCATCCACGAGAGTGAACATATATGTTGCCATC 746
Qy      221 AsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysLysThrHisArg 240
Db      747 AACCTGATCGTGAAGCAATCCAGAGCATTTCTGAATGTGACATCTGCAAATGACACGA 806
Qy      241 GlyGlySerAsnGlyArg 246
Db      807 GGAGGATCCCATGGCGG 824

RESULT 3
LOCUS   BM467984 1118 bp mRNA linear EST 05-FEB-2002
DEFINITION AGENCOURT 6437937 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:5552885
5', mRNA sequence.
ACCESSION BM467984
VERSION   BM467984.1 GI:18517026
KEYWORDS EST.
SOURCE    Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS   NIH-MGC http://mgi.nci.nih.gov/.
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished
COMMENT   Contact: Robert Strausberg, Ph.D.
          Email: cgabs-r@mail.nih.gov
          Tissue Procurement: ATCC
          cDNA Library Preparation: Life Technologies, Inc.
          DNA Sequencing by: The I.M.A.G.E. Consortium (LNLML)
          DNA Sequencing by: Agencourt Bioscience Corporation

```

```

Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: LHAM12216 row: n column: 22
High quality sequence stop: 689.
Location/Qualifiers
1..1118
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5532885"
/tissue_type="leiomyosarcoma"
/lab_host="PH10B (phage-resistant)"
/clone_11b="NIH_MGC_71"
/notes="Organ: uterus; Vector: pCMV-SPORT6; Site_1: Not1,
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2.1 kb."
BASE COUNT 234 a 311 c 361 g 209 t 3 others
ORIGIN
Alignment Scores:
Pred. No.: 6,47e-220 Length: 1118
Score: 230.00 Matches: 259
Percent Similarity: 99.23% Conservative: 0
Best Local Similarity: 99.23% Mismatches: 1
Query Match: 83.03% Indels: 2
DB: 12 Gaps: 0
US-09-896-522-2 (1-277) x BM467984 (1-1118)
Qy      1 MetAlaSerAlaGlyGlyGlnAspCysGlnSerProAlaProGlnAlaAspArgProHis 20
Db      47 ATGCGCTTCGGCGGAGCGAAGACTGCGAGAGCCCGCGGAGCCGACCGTCCGCAC 106
Qy      21 GlnArgProPheLeuIleGlyValSerGlyGlyThrAlaSerGlyLysSerThrValCys 40
Db      107 CAGCGGCCCTTCTGATAGGGGTGAGCGCGGACATGCCAGCGGGAAGTGCACCGTGTG 166
Qy      41 GlnLysIleMetGlnLeuLeuGlnGlnAlaGlnGlnValGlnGlnArgGlnArgLysValVal 60
Db      167 GAGAGATCATGAGTGTCTGAGGACAGAACGAGTGAACAGCGCGCAGCGGAGGAGTGTG 226
Qy      61 IleLeuSerGlnAspArgPheTyrlsValLeuThrAlaGlnGlnLysAlaLysAlaLeu 80
Db      227 ATCTGAGCCAGAGCAGGTTCTTCAAGGTCTGAGCGCAGACGAGAGGCGAAGGCTTG 286
Qy      81 LysGlyGlnTyrlsAspPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThr 100
Db      287 AAGGACATTCATCAATTTTGAACATCCAGATGCTTGTGAATGATTTGATGACAGGACT 346
Qy      101 LeuLysAsnIleValGlnGlyLysThrValGlnValProThrTyrlsAspPheValThrHis 120
Db      347 CTGAGAGACATCATGTGAGGCGCAAAACGAGTGAAGTCCGACCTATGATTTGTGACACAC 406
Qy      121 SerArgLeuProGlnThrThrValValTyrlsProAlaAspValValLeuPheGlnGlyIle 140
Db      407 TCAAGGTACCAAGACCAAGGTGTCTACCTGCGGAGTGTGTTGTTAGGGGCATC 466
Qy      141 LeuValPheTyrlsSerGlnGlnIleArgAspMetPheHisLeuArgLeuPheValAspThr 160
Db      467 TTGGTGTCTTACAGCAGAGAGATCCGGGACATGTTCCATCTGCGCTTCTGTGAGACAC 526
Qy      161 AspSerAspValArgLeuSerArgArgValLeuArgAspValArgArgGlyArgAspLeu 180
Db      527 GACTCGAGTCAGAGGTGTCTGGAAGAGTTCGCGGAGCGTCCGCGGAGGAGGAGCCTG 586
Qy      181 GlnGlnIleLeuThrGlnTyrlsThrThrPheValLysProAlaPheGlnGlnGlyLeu 200
Db      587 GAGCAGATTCAGCAGAGTACACACCTTCGTGAAGCGGCTTGAAGAGTTCGCTGTG 646
Qy      201 ProThrLysLysTyrlsAlaAspValIleIleProArgGlyValAspAspMetValAlaIle 220
Db      647 CCGACAAAGAAATGAGCGATGTGATCATCCACGAGAGGTGACATATATGTTGCCATC 706

```

QY 221 AsnleuileValGlnHisIleGlnAspIle-LeuAsnGlyAspIleCysValSTpHisAR 240
 DB 707 AACCTGATCGTGCAGCATTCAGACATTTCTGAT-GGTGACATCTGCAGATGCGACCG 765
 QY 240 GGLGlySerAsnGlyTrpSerGlyTrpArgThrPheSerGlyProGlyAspHisProGly 260
 DB 766 AGAGGGGTCCATGGCGGAGCTCAACCGGACCTTTTCTGAGCCAGGGAGCCACCTCG 825
 QY 260 Y 260
 DB 826 N 826
 RESULT 4
 CD514811 898 bp mRNA linear EST 06-JUN-2003
 LOCUS AGNCCOURT 14376196 NIH MGC 181 Homo sapiens cDNA clone
 DEFINITION IMAGE:30396201 5', mRNA sequence.
 CD514811
 VERSION CD514811.1 GI:31446529
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 898)
 AUTHORS NIH-MGC http://mgs.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished
 COMMENT Contact: Daniela S. Gerhard, Ph.D.
 Office of Cancer Genomics
 National Cancer Institute / NIH
 Bldg. 31 Rm10A07 Bethesda, MD 20892
 Email: cgabs-r@mail.nih.gov
 Tissue Procurement: Dr. Michael Brownstein
 cDNA Library Preparation: Invitrogen Corp
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: NDAM477 row: i column: 10
 High quality sequence stop: 662.
 FEATURES
 source location/Qualifiers
 1..898
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:30396201"
 /tissue_type="White Matter"
 /dev_stage="Unknown"
 /lab_host="DH10B-Ton A (T1 and T5 phage resistances)"
 /clone_id="NIH_MGC_181"
 /note="Vector: PCMV-SPORT6.1; Site 1: NotI; Site 2: EcoRV
 (destroyed); Library is oligo-dT primed and directionally
 cloned (EcoRV site is destroyed upon cloning). Average
 insert size 1.42 kb. Library was constructed by
 (Invitrogen). Note: this is a NIH_MGC Library."
 BASE COUNT 194 a 231 c 310 g 158 t 5 others
 ORIGIN
 Alignment Scores:
 Pred. No.: 5.18e-218 Length: 898
 Score: 228.00 Matches: 228
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 82.31% Indels: 0
 DB: 14 Gaps: 0
 US-09-896-522-2 (1-277) x CD514811 (1-898)
 QY 1 MetAlaSerAlaGlyGlyGlyAspCysGlySerProAlaProGlyAlaAspArgProHis 20
 |||||

DB 72 ATGGCTTCGCGGAGGCGAAGACTGCGAGAGCCCGCGGAGGCGCAGCTCCGCAC 131
 QY 21 GlnArgProPheLeuIleGlyValSerGlyGlyThrAlaSerGlyLysSerThrValCys 40
 DB 132 CACGGCGCCCTCTCTGATGAGGGGTGACCGCGGCGACCGCAGCGGAAGTCGACCGGTGT 191
 QY 41 GlnLysIleWecGlyLeuLeuGlyGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 60
 DB 192 GAGAGAGTCAATGAGTGTCTGGAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 251
 QY 11LeuSerGlnAspArgPheTrpLysValLeuThrAlaGlnGlnLysAlaLysAlaLeu 80
 DB 252 ATCCTGAGCAGAGACAGGTTCTACAGGTCTCTACGCGCAGAGAGAGAGAGAGAGAGAG 311
 QY 81 LysGlyGlnTrpAsnPheAspHisProAlaPheAspAsnAspLeuMetHisArgThr 100
 DB 312 AAGGACAGTACATTTTGCATTCACAGATGCTTGTGATATGATTTGTGACAGAGACT 371
 QY 101 LeuLysAsnIleValGlnGlyLysThrValAlaGlnValProThrTrpAspPheValThrHis 120
 DB 372 CTGAGAGACATCGTGAAGGAGCAAGCGTGAAGGTCCGACCTATGATTTTGTGACACAC 431
 QY 121 SerArgLeuProGlyThrThrValValTyrProAlaAspValLeuPheGlnGlyIle 140
 DB 432 TCAAGGTACCAAGACACAGCGGTGTCTACCTCGCGAGCGTGTCTGTTGAGGCGATC 491
 QY 141 LeuValPheTrpSerGlnGlnIleArgAspMetPheHisLeuArgLeuPheValAspThr 160
 DB 492 TTGGTGTTCACAGCGAGAGATCCGAGCATGTCACCTGCGCTCTTCGCGACAC 551
 QY 161 AspSerAspValArgLeuSerArgValLeuArgAspValArgArgGlyArgAspLeu 180
 DB 552 GACTCCGACGTCGAGCTGTCTCAAGAGTCTCCGAGAGCTGCGCGAGAGAGAGAGCTG 611
 QY 181 GlnGlnIleLeuThrGlnTrpThrPheValLysProAlaPheGlnGlnGlyLeu 200
 DB 612 GAGCAGATTCGTGACGAGTACACACCTTCGTGAACCGGCTTCGAGGAGTCTCGCTG 671
 QY 201 ProThrLysLysTrpAlaAspValIleIleProArgGlyValAspAsnMetValAlaIle 220
 DB 672 CCGACAAAGAGATGCCGATGATCATCCACGAGAGTGAACAATATGTTGTCATC 731
 QY 221 AsnLeuIleValGlnHisIleGln 228
 DB 732 AACCTGATCGTGCAGCATTCGAG 755
 RESULT 5
 BI256928 725 bp mRNA linear EST 17-JUL-2001
 LOCUS 602975534F1 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:5114816 5',
 DEFINITION mRNA sequence.
 BI256928
 VERSION BI256928.1 GI:14811813
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 725)
 AUTHORS NIH-MGC http://mgs.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabs-r@mail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: Incyte Genomics, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: LHAM1279 row: c column: 09
 High quality sequence stop: 707.


```

FEATURES
    source
        location/Qualifiers
            1..725
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="IMAGE:5114816"
                /tissue_type="cervical carcinoma cell line"
                /lab_host="DH10B"
                /clone_lib="NIH_MGC_12"
                /note="Organ: cervix; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.4 kb. Library prepared by Life Technologies."
BASE COUNT      175 a      181 c      224 g      145 t
ORIGIN
Alignment Scores:
Pred. No.:      4,19e-215      Length:      725
Score:          225.00      Matches:      225
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    81.23%      Indels:      0
                Gaps:      0
US-09-896-522-2 (1-277) x B1256928 (1-725)
QY      36  LysSerThrValCysGluValIleMetGluLeuGluGlnAaGlnValGluGlnArg 55
Db      3  AAGTCGACCGGTGTGAGAGATCATGAGTTGCTGGGACAGAACGAGTGGAAACAGCGG 62
QY      56  GlnArgValValIleLeuSerGlnAspArgPheTyrValLeuThrAlaGluGln 75
Db      63  CAGCGAAGGTTGATCTGATCTGAGCCAGAGAGGTTCTCAAGGTCCTGACGCGACAGCAG 122
QY      76  LysAlaLysAlaLeuLysGluGlnTyrAsnPheAspHisProAspAlaPheAspAsnAsp 95
Db      123  AAGGCCAAGGCTTGAAGAGACATGACATTTTACCATCCAGATGCTTGAATGAT 182
QY      96  LeuMetHisArgThrLeuLysAsnIleValGluGluTyrValGluValProThrTyr 115
Db      183  TTGATGACACGAGACTCTGAGAACATCGTGAAGGCGAAACGGTGAAGTGCCTGACCTAT 242
QY      116  AspPheValThrHisSerArgLeuProGluThrThrValValTyrProAlaAspValVal 135
Db      243  GATTTTGACACACTAGAGGTACCAAGACCAAGGTTGTTACCTCGGAGACGTGAT 302
QY      136  LeuPheGluGluIleLeuValPheTyrSerGlnGluIleArgAspMetPheHisLeuArg 155
Db      303  CTGTTTGAGGGCATCTTGATGTTCTACAGCCAGAGAGATCCGGGACATGTTCCACTGGC 362
QY      156  LeuPheValAspThrAspSerAspValArgLeuSerArgArgValLeuArgAspValArg 175
Db      363  CTCTTCGTGACACGAGCTCGAGAGTCAAGCTGTCTGAAGAGTTCCTCCGGGACGTCCG 422
QY      176  ArgGluArgAspLeuGluGlnIleLeuThrGlnTyrThrPheValLysProAlaPhe 195
Db      423  CGAGGGAGGGACCTCGAGACGATTTCTGACGACATCCACTTCTGTGAAGCCGGCTTC 482
QY      196  GlnGluPheCysLeuProThrLysLeuTyrAlaAspValIleIleProArgGluValAsp 215
Db      483  GAGGAGTTTCGCTGCGCCGACAAAGAGATATCCGATGTGATCCACAGAGAGTGGAC 542
QY      216  AsnMetValAlaIleAsnLeuIleValGlnHisIleGlnAspIleLeuAsnGluAspIle 235
Db      543  AATATGTTGTCATCAACCTGATCGTGCACACATCCAGGACATTCGATGTGATGCATC 602
QY      236  CysLeuTrpHisAspGluGluSerAsnGluValArgSerTyrLysArgThrPheSerGluPro 255
Db      603  TGCAAATGGACCCAGAGAGGATCCAATGGCGGAGCTACAAAGCGAATTTTCTTGAGCA 662
QY      256  G1yAspHisProGly 260
Db      663  GGGGACCACTGGG 677

```

```

RESULT 6
BG826894
LOCUS      916 bp      mRNA      linear      EST 22-MAY-2001
DEFINITION 602750978F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:4903819 5',
mRNA sequence.
ACCESSION  BG826894
VERSION    BG826894.1 GI:14174481
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 916)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished
COMMENT    Contact: Robert Strausberg, Ph.D.
           Email: cga@bbs-romail.nih.gov
           Tissue Procurement: ATCC
           cDNA Library Preparation: Ling Hong/Rubin Laboratory
           cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
           DNA Sequencing by: Incyte Genomics, Inc.
           Clone distribution: MGC clone distribution information can be
           found through the I.M.A.G.E. Consortium/LNL at:
           http://image.lnl.gov
           Plate: LNCMI802 row: k column: 20
           High quality sequence stop: 843.
FEATURES
    source
        location/Qualifiers
            1..916
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="IMAGE:4903819"
                /tissue_type="rhabdomyosarcoma"
                /lab_host="DH10B (phage-resistant)"
                /clone_lib="NIH_MGC_17"
                /note="Organ: muscle; Vector: pORF7; Site_1: EcoRI; Site_2: XhoI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAAGG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT      203 a      239 c      309 g      165 t
ORIGIN
Alignment Scores:
Pred. No.:      5.73e-206      Length:      916
Score:          216.00      Matches:      216
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    77.98%      Indels:      0
                Gaps:      0
US-09-896-522-2 (1-277) x BG826894 (1-916)
QY      1  MetAlaSerAlaGlyGluAspCysGluSerProAlaProGluAlaAspArgProHis 20
Db      80  ATGGCTTCGGCGGAGGAGGCAAGACTCGAGAGCCCGCGCGGAGGCCAGCTCCGGAC 139
QY      21  GlnArgProPheLeuIleGlyValSerGlyGlyThrAlaSerGlyLysSerThrValCys 40
Db      140  CACCGGCCCCCTCTGATAGGGGTGACCGGCGCACATCCAGCGGAAAGTCGACCGGTGT 199
QY      41  GlnLysIleMetGluLeuLeuGluGlnAaGlnValGluGlnArgGlnArgLysValVal 60
Db      200  GAGAAGATCTAGAGTTGCTGGGACAAACAGAGTGAACAGCGGACGGAAGGTGGTC 259
QY      61  IleLeuSerGlnAspArgPheTyrLysValLeuThrAlaGluGlnLysAlaLysValLeu 80
Db      260  ATCTTAGGACGAGACAGTTCTCAAGAGTCTTACGCGCAGAGACGAAGGCGAAGCCTTG 319

```

```

Qy      81  LysGIyGlnTYrAsnPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThr 100
Db      320 AAAGACAGTCAATTTGGACCATCCAGATGCCCTTGTATATGATTTGATCAGACGACT 379
Qy      101  LeuLYsAsnIleValGIuGIyLYsThrValGIuValProThrTYrAspPheValThrHis 120
Db      380 CTGAAGAATCATCTGTGAGGGCAAAACGGTGGAGGTGCGCATATGATTTGTGACACAC 439
Qy      121  SerArgLeuProGIuThrThrValValTYrProAlaAspValIleLeuPheGIuGIyIle 140
Db      440 TCAAGGTTCACAGAACACACGGTGTCTACCCCTGGACGGTGTCTGTTTGAGGGCACTC 499
Qy      141  LeuValPheTYrSerGIuGIuIleArgAspMetPheHisLeuArgLeuPheValAspThr 160
Db      500 TTGGTGTTCTACAGCCAGACGATCCGGACATGTTCCACCTGCCCTCTTGTGTGACACC 559
Qy      161  AspSerAspValArgLeuSerArgArgValLeuArgAspValArgArgGIyArgAspLeu 180
Db      560 GACTCCGACGTACGGCTGTCTCGAAGAGTTCTCCGGACGTGCGCCGAGGAGGACCTG 619
Qy      181  GluGlnIleLeuThrGlnTYrThrThrPheValLYsProAlaPheGIuGIuPheCysLeu 200
Db      620 GAGCAGATTCTGACCGACATACACACCTTCGTGAAGCCGCGCTTCGAGGAGTTCGCTG 679
Qy      201  ProThrLYsLYsTYrAlaAspValIleIleProArgGIyValAspAsn 216
Db      680 CCGAACAAAGATATGCCGATGTATCATCCACGAGGAGTGACAAAT 727

RESULT 7
LOCUS   BQ072501 1036 bp mRNA linear EST 02-APR-2002
DEFINITION AGENCOURT 6838909 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:5761694
5', mRNA sequence.
ACCESSION BQ072501
VERSION   BQ072501.1 GI:19901547
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 1036)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished
COMMENT    Contact: Robert Strausberg, Ph.D.
           Email: cgs@bbs-remail.nih.gov
           Tissue Procurement: Life Technologies, Inc.
           cDNA Library Preparation: Life Technologies, Inc.
           DNA Sequencing by: Agencourt Bioscience Corporation
           Clone distribution: MGC clone distribution information can be
           found through the I.M.A.G.E. Consortium/LNLN at:
           http://image.jlm.gov
           Plate: LLM12810 row: 1 column: 15
           High quality sequence stop: 644.
           Location/Qualifiers
             1..1036
               /organism="Homo sapiens"
               /mol_type="mRNA"
               /db_xref="taxon:9606"
               /clone="IMAGE:5761694"
               /lab_host="DH10B"
               /clone_1ib="NIH MGC 122"
               /note="Organ: pooled lung and spleen; Vector: pCMV-SPORT6;
               Site_1: NoCI; Site_2: ECoRV (destroyed); RNA source
               anonymous pool of 24 week female lung, 16 week female
               spleen, and 20-22 week male spleens. Library is oligo-dT
               primed and directionally cloned (ECoRV site is destroyed
               upon cloning). Average insert size 1.4 kb. Insert size
               range 1-3 kb. Library is normalized and enriched for
               full-length clones and was constructed by C. Gruber
               (Invitrogen). Research Genetics tracking code 026. Note:

```

```

BASE COUNT 221 a 285 c 346 g 184 t
ORIGIN
Alignment Scores:
Pred. No.: 6.63e-204 Length: 1036
Score: 214.00 Matches: 227
Percent Similarity: 99.56% Conservative: 0
Best local Similarity: 99.56% Mismatches: 0
Query Match: 77.26% Indels: 1
DB: 13 Gaps: 0

US-09-896-522-2 (1-277) x BQ072501 (1-1036)
Qy      1  MetAlaSerAlaGIyGIuAspCysGlnSerProAlaProGIuAlaAspArgProHis 20
Db      64 ATGGCTTGGCGGAGGCGAAGACTGCGAAGCCCCGCGGAGGCCGACCTCCGAC 123
Qy      21  GlnArgProPheLeuIleGIyValIleSerGIyTYrAlaSerGIyLYsSerThrValCys 40
Db      124 CAGGGGCCCTTCTGATAGGGGTGAGGGCGGCACTGCCAGCGGGAAGTGCAGCGTGT 183
Qy      41  GluLYsIleMetGluLeuLeuGIyGlnAsnGluValGIuGlnArgGlnArgLYsVal 60
Db      184 GAGAGATCATGTGAGTGTCTGGGACAGAACGAGTGGAAACGCGCAGCGGAGTGTG 243
Qy      61  IleLeuSerGlnAspArgPheTYrLYsValLeuThrAlaGIuGlnLYsAlaLYsAlaLeu 80
Db      244 ATCTGAGCCAGACACAGTCTTCAGAGTCTTACCGGACGAGACGACGACGCTTG 303
Qy      81  LysGIyGlnTYrAsnPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThr 100
Db      304 AAAGACAGTACATTTTGGACCATCCAGATGCTTGATATATGATTTGATGACAGACT 363
Qy      101  LeuLYsAsnIleValGIuGIyLYsThrValGIuValProThrTYrAspPheValThrHis 120
Db      364 CTGAAGAATCATCTGTGAGGGCAAAACGGTGGAGTCCGACCTATGATTTGTGACACAC 423
Qy      121  SerArgLeuProGIuThrThrValValTYrProAlaAspValIleLeuPheGIuGIyIle 140
Db      424 TCAAGGTTCACAGAACACACGGTGTCTTACCTCGGACGTGTTCTGTTTGAGGGCATC 483
Qy      141  LeuValPheTYrSerGIuGIuIleArgAspMetPheHisLeuArgLeuPheValAspThr 160
Db      484 TTGGTGTTCTACAGCCAGACGATCCGGACATGTTCCACTGCGCTCTCGTGACACC 543
Qy      161  AspSerAspValArgLeuSerArgArgValLeuArgAspValArgArgGIyArgAspLeu 180
Db      544 GACTCCGACGTACGGCTGTCTCGAAGAGTTCCTCCGGACGTGCGCGGAGGAGGACCTG 603
Qy      181  GluGlnIleLeuThrGlnTYrThrThrPheValLYsProAlaPheGIuGIuPheCysLeu 200
Db      604 GAGCAGATTCTGACCGACGACACACCTTCGTGAACCGGCTTCAGAGAGTTCCTCG 663
Qy      201  -ProThrLYsLYsTYrAlaAspValIleIleProArgGIyValAspAsnMetValAlaI 220
Db      664 CCGACAAAGATATGCCGATGTGATCATCCACGAGGAGTGGACAAATATGTTGCCAT 723
Qy      220  eAnIleuIleValGIuHisIle 227
Db      724 CAACCTGATGTGACACACATC 745

RESULT 8
LOCUS   BG770518 730 bp mRNA linear EST 15-MAY-2001
DEFINITION 60274236F1 NIH_MGC_49 Homo sapiens cDNA clone IMAGE:4859577 5',
mRNA sequence.
ACCESSION BG770518
VERSION   BG770518.1 GI:14081171
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

REFERENCE
1 (bases 1 to 730)
AUTHORS
TITLE
JOURNAL
COMMENT
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
NIH-MGC <http://mgs.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC/DCTD/DRP
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
<http://image.llnl.gov>
Plate: L1CM1715 row: h column: 10
High quality sequence stop: 720.
Location/Qualifiers
1..730
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4859577"
/tissue_type="melanotic melanoma, high MDR (cell line)"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 49"
/note="Organ: skin; Vector: pOTB7, site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH-MGC
Library."

BASE COUNT 156 a 190 c 258 g 126 t
ORIGIN

Alignment Scores:
Pred. No.: 4,576-203 Length: 730
Score: 213.00 Matches: 213
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 76.90% Indels: 0
Gaps: 0

US-09-896-522-2 (1-277) x BG770518 (1-730)

QY 1 MetAaSerAlaGlyGluAspCysGluSerProAlaProGluAlaAspArgProHis 20
Db 82 ATGCGCTCGGGGAGGAGGAGAGCTGCGAGAGCCCGCGCGAGCCGACCGCCAC 141
QY 21 GlnArgProPheLeuIleGlyValSerGlyValThrAlaSerGlyLysSerThrValCys 40
Db 142 CAGCGGCCCTTCCTCGATAGGGGTGAGCGCGCGCACTCCAGCGGGAAGTGCACCGTGT 201
QY 41 GluValIleMetGluLeuGluValGlnAsnGluValGluGlnArgGlnArgGlyValVal 60
Db 202 GAGAGAGTCATGAGGTGCGGACAGAAACAGAGTGAACCGCGCAGCGGAAGGTGCTC 261
QY 61 IleLeuSerGlnAspArgPheTyrGlyValLeuThrAlaGluGlnLysAlaValLeu 80
Db 262 ATCTGAGCCAGGACAGGATCTTCAAGTCTCTGACGCGAGAGCAGAAAGCCAGCTTG 321
QY 81 LysGlyGlnTyrAsnPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThr 100
Db 322 AAAGAGACAGTACATTTTGAACCATCCAGATGCTTTATATATGATTGATGACAGACT 381
QY 101 LeuValSerIleValGluGlyValThrValGluValProThrTyrAspPheValThrHis 120
Db 382 CTGAGAGACATCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 441
QY 121 SerArgLeuProGluThrThrValValTyrProAlaAspValValLeuPheGluGlyIle 140

Db 442 TCAAGTTACAGAGACACAGGTGTCTACCTCGCGAGCGTGGTCTGTTCGAGGCATC 501
QY 141 LeuValPheTyrSerGlnGluIleArgAspMetPheHisLeuArgLeuPheValAspThr 160
Db 502 TTGGTGTTCACAGCCAGAGGATCCGGACATGTCTCCAGCTGCGCTTCTGTGACACC 561
QY 161 AspSerAspValArgLeuSerArgValLeuAlaAspValArgArgGlyArgAspLeu 180
Db 562 GACTCCGACATCAGGCTGTCTCCAGAGTCTCCGGAGCTGCGCGAGGAGGAGGAGCTG 621
QY 181 GluGlnIleLeuThrGlnTyrThrPheValLysProAlaPheGluGlnPheCysLeu 200
Db 622 GACAGATTCGTGACGACATCACACCTTGTAACCGGCGCTTCAGAGAGTCTGCTG 681
QY 201 ProThrLysLysTyrAlaAspValIleIleProArgGly 213
Db 682 CCGACAAAGATGATGCCGATGATCATCCACAGAGA 720

RESULT 9
BG491384
LOCUS
DEFINITION
602535670F1 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:484936 5',
mRNA sequence.
BG491384
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 869)
NIH-MGC <http://mgs.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: DCTD/DRP
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
<http://image.llnl.gov>
Plate: L1CM1493 row: k column: 17
High quality sequence stop: 863.
Location/Qualifiers
1..869
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:484936"
/tissue_type="melanotic melanoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 41"
/note="Organ: skin; Vector: pOTB7, site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH-MGC Library."

BASE COUNT 197 a 223 c 291 g 156 t 2 others
ORIGIN

Alignment Scores:
Pred. No.: 5,526-203 Length: 869
Score: 213.00 Matches: 239
Percent Similarity: 99.17% Conservative: 0
Best Local Similarity: 99.17% Mismatches: 0
Query Match: 76.90% Indels: 2
Gaps: 0

US-09-896-522-2 (1-277) x BG491384 (1-869)

QY	1	MetAlaSerAlaGlyGlyGluAspCysGluSerProAlaProGluAlaAspArgProHis	20
Db	52	ATGGCTTGGCGGAGGCGAACAACGTCCAGAGCCCCCGCGGAGCGGACGCTCCGAC	111
QY	21	GlnArgProPheLeuIleGlyAlaSerGlyGlyThrAlaSerGlyLysSerThrValCys	40
Db	112	CAGCGGCCCTTCTCATATAGGGGTGACGGCGGCACATGCCAGCGGGAAGTCCACTGTGT	171
QY	41	GlyLysIleMetGluLeuLeuGlyGlnAsnGluValGluGlnArgGlnArgValVal	60
Db	172	GAGAAAGATCATGGAATTGCTGGGACAGAACGAGTGGAAACGCGCACGGAAGGTGTCT	231
QY	61	IleLeuSerGlnAspArgPheThrLysValLeuThrAlaGluGlnLysAlaLysAlaLeu	80
Db	232	ATCTCGAACCAGAGACAGGTTCTCAACAGTCTCTGACGCGACAGACAAAGGCCAAGCCTTG	291
QY	81	LysGlyGlnTyrAsnPheAspHisLeuProAlaAlaPheAspAsnAspLeuMetHisArgThr	100
Db	292	AAAGGACAGTCAATTTTGACCATCCAGATGCCCTTGATTAATGATTTATGACACAGAGACT	351
QY	101	LeuLysAsnIleValGluGlyLysThrValGluValProThrTyrAspPheAlaThrHis	120
Db	332	CTGAAGAACATCGTAGAGGCAAAACGGTGGAGGCGGACCTATGATTTTGTGTGACAC	411
QY	121	SerArgLeuProGluThrThrValValTyrProAlaAspValValLeuPheGluGlyIle	140
Db	412	TCAAGGTTCACGAGACACACGGTGTCTTACCTCGCGAGCGTGGTCTGTGGAGGGCANTC	471
QY	141	LeuValAlaPheTyrSerGlnGluIleArgAspMetPheHisLeuArgLeuPheValAspThr	160
Db	472	TTGGTGTTCTAAGCCAGAGAGATCCGGGACATGTTCACCTGCGCTCTTCTGTGACACC	531
QY	161	AspSerAspValArgLeuSerArgArgValLeuArgAspValArgArgGlyArgAspLeu	180
Db	532	GACTCCGACGTCAGGCTGCTCTGGAAGAGTTCCTCGGAGCGTGCCTCGAGGGAGGACCTG	591
QY	181	GluGlnIleLeuThrGlnTyrThrThrPheValLysProAlaPheGluGluPheLysLeu	200
Db	592	GAGCAGATTTCGACCCAGTACACCACTTCGGAAGCGCGGCTTCGAGAGAGTTCTGCTCG	651
QY	201	ProThrLys-LysTyrAlaAspValIleIleProArgGlyValAlaAspAsnMet-ValAlaI	220
Db	652	CCGACAAAGAAGTATGCCGATGTGATCATCCACAGAGAGTGGACAAATATAGGTTGCCA	711
QY	220	IleAsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyLysAspIleCysLysTrpHis	239
Db	712	TCAACTTATCGTGGACACATCCAGGACATTTCTGAATGGTGTGACATTCGCAATGGCAC	770

REFERENCE 1 (bases 1 to 956)
AUTHORS NIH-MGC <http://mgc.ncl.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strusberg, Ph.D.
Email: cgs@bcr-remail.nih.gov
Tissue Procurement: Dr. James R. Lupski
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LNL at
http://image.lnl.gov
plate: L1AM13574 row: b column: 18
High quality sequence stop: 552.

FEATURES

Source

1. 956

/organism="Homo sapiens"

```
/mol_type="mRNA"
```

/db_xref="taxon:9606"

```
/clone="IMAGE:6184625"
```

```
/sex="male"
```

```

/tissue_type="dorsal root ganglia"
/dor_at3300="adult 36 wks"

```

```
/dev_stage="adult, 36
/lab_host="DPU0B"
```

```
/Lab_host="DH10B"
```

```
/clone_lib="Lupski_dorsal_foot_ganglion"
/notes="vector: nCMV-SPORT6 (Life Technol
```

```

/notes=Vector: pcmv-sporib (Life Technologies); site_1:
NOT: Site 2: Salt: cDNA made by oligo-dT priming.

```

Directionally cloned using the following adaptors:
NotI; site 2: salI; cDNA made by oligo-a1 priming.

directionally cloned using the following adapters: 5'-TCGACCCACGGCTCG-3' and

5'-GACTAGTCTAGATCGCGAGCGGCGCCCT(15)-3'. Size selected

1 kb for average insert length 1.7 kb. This is a primary

library, non-amplified. Library constructed by Life

Technologies and donated by J. Lupski, M.D./Ph.D. (Bay)

College of Medicine) and is available through Life

Technologies."

Technologie

BASE COUNT	202 a	272 c	296 g	186 t
ORIGIN				

ORIGIN

Alignment Scores:

Assignment scores:	
Pred. No.:	6.12e-203
Length:	956

Score:	213.00	Matches:	213
--------	--------	----------	-----

Percent Similarity:	100.00%	Conservative:	0
---------------------	---------	---------------	---

```
Best Local Similarity: 100.00% Mismatches: 0
```

Query Match:	76.90%	Indels:	0
--------------	--------	---------	---

DB:	13	Gaps:	0
-----	----	-------	---

US-09-896-522-2 (1-277) X BQ719741 (1-956)

Qy	6	GLYGLNAsPCysGluSerProAlaProGluAlaAspAlaProHisLeuSerGlnAspProPheLeu	23
Db	22	GGGGAAGACTGCAGAGAGCCCGCGCGGAGCGGACCGTCCGCACCAAGCGCCCTTCTCGT	81
Qy	26	ILGLIValSerGlyGlyThrAlaSerGlyLysSerThrValCysGluLysILlMetGlu	45
Db	82	ATAGGGGTGAGCGCGGCACCTCCAGCGGGAAGTGACCGTGTGTGAAGAATCATTTGAG	143
Qy	46	LeuLeuGlyGlnAsnGluValGluGlnArgGlnArgLysValIleIleLeuSerGlnAsp	65
Db	142	TTGCTGGGACAGAACAGAGGTGAGACAGCGGACCGGAAAGTGTGATCTTACACCAAGAC	203
Qy	66	ArgPheTyrLysValLeuThrLagluGlnLysAlaLysAlaLeuLysGlyGlnTyrAsn	85
Db	202	AGGTTCTACAAAGTCTCTGACGGCAGAGCAAGAGCCAAAGCCTTGAAAGGACAGTACAT	263
Qy	86	PheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThrLeuLysAsnIleVal	105
Db	262	TTTGACCAATCCAGATGCTCTTGTATATGATTTGATGACACAGAGACTCTGAAAGAACATCGTG	323
Qy	106	GluGlyLysPheThrValGluValProThrTyrAspPheValIleHisSerArgLeuProGlu	125
Db	322	GAGGGCAAAAACGGTGGAGGTGCCGACTATGATTTTGTGACACACTCAAGGTATCCAGAG	383
Qy	126	ThrThrValValTyrProAlaAspValIleLeuPheGluGlyIleLeuValPheTyrSer	145
Db	382	ACCAACGGTGGTCTACCCCTGCGGACGTGGTCTGTTGAGGGGCATCTGGTCTTACAGC	443
Qy	146	GlnGluIleArgAspMetPheHisLeuArgLeuPheValAspThrAspSerAspValArg	165
Db	442	CAGGAATCCGGGACATGTTCCACCTGGGCGCTTCGTGGACACGCACTCCGACGTGAGG	503
Qy	166	LeuSerArgArgValLeuArgAspValArgArgGlyArgAspLeuGluGlnIleLeuThr	185
Db	502	CTGTCTTCGAAGAGTTCTCCGGGACGTGGCCGACGGAAGGAGACTGGAGACAAATTTCTGCG	563

QY 186 GltYrThrThrPheValIysProAlaPheGluGluPheCysLeuProThrIlyLeuTyr 205
 DB 562 CAGTACCCACCTTCGTGAAGCCGGCCTTCGAGGAGTTCTGCTCGCAGCAAGAGTAT 621
 QY 206 AlaSPValIleIleProArgGlyValAspAsnMetVal 218
 DB 622 GCCGATGTGATCATCCCGCAGAGTGGACATATGTGT 660
 RESULT 11
 BG390519 989 bp mRNA linear EST 12-MAR-2001
 LOCUS 60241618.F1 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:4524308 5',
 DEFINITION mRNA sequence.
 ACCESSION BG390519 GI:13283967
 VERSION BG390519.1
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE NIH-MGC http://mgs.nci.nih.gov/.
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 TITLE Unpublished
 JOURNAL
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: LHAM10428 row: F column: 21
 High quality sequence stop: 750.
 Location/Qualifiers
 1..989
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4524308"
 /tissue_type="embryonal carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_92"
 /note="Organ: testis; Vector: pCMV-SPORT6; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally; oligo-dt primed.
 Average insert size 2.5 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH_MGC library."
 BASE COUNT 256 a 258 c 314 g 161 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 6.56e-198 Length: 989
 Score: 208.00 Matches: 208
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 75.09% Indels: 0
 DB: 10 Gaps: 0
 US-09-896-522-2 (1-277) x BG390519 (1-989)
 QY 7 GluAspCysGluSerProAlaProGluAlaAspArgProHisGlnArgProPheLeuIle 26
 DB 2 GAAGACTGCAGAGCCCCCGCGCGAGCGACCGTCGACACGAGCGCCCTTCCTGATA 61
 QY 27 GlyValSerGlyGlyThrAlaSerGlyIleSerThrValCysGluValIleMetGluLeu 46
 DB 62 GGGGTGAGCGCGGCGACTGCGAGCGAGAGTTCAGCCGTGTGTGAAGATCATGAGATTG 121
 QY 47 LeuGlyGlnAsnGluValGluGlnArgGlnArgValValIleLeuSerGlnAspArg 66

DB 122 CTGGAGCAGAACGAGGTGAACAGCGCAGCGAAGGTGTCATCTCGACCGACGACAG 181
 QY 67 PheTyIlysValIleuThrAlaGluGlnIlyValAlaLeuIlyseGlyGlnTyraenPhe 86
 DB 182 TTCACAGAGTCTCTGACGCGAGAGAGAGAGCCAGGCCCTTGAAAGACAGTCAATTTT 241
 QY 87 AspHisProAspAlaPheAspAsnAspLeuMetHisArgThrIleuIysAsnIleValGlu 106
 DB 242 GACCATCCAGATCCCTTTGATTAATGATTGATCACAAGACTCTGAAACATCTGGAG 301
 QY 107 GlyIysThrValGluValProThrTyraSPheValThrHisSerArgLeuProGluThr 126
 DB 302 GCGAAGACGGTGAAGGCGGACCTTAATGATTTGTGACACACTCAAGGTTACAGAGACC 361
 QY 127 ThrValValTytrProAlaAspValIleuPheGluGlyIlyleuValPheTyrsSerGln 146
 DB 362 ACGGTGTCACCTCGCGAGCGTGTCTGTGAGAGGACATCTGTGTCTACACCCAG 421
 QY 147 GluIleArgAspMetPheHisIleuArgLeuPheValAspThrAspSerAspValArgLeu 166
 DB 422 GAGATCCGGACATGTTCCTCACTGCCCTCTTGTGACACGACCTCGACGTCAAGCTG 481
 QY 167 SerArgArgValIleuArgAspValArgArgGlyArgAspLeuGluGlnIleuThrGln 186
 DB 482 TCTCGAAGAGTTCTCCGGAGCGTGGCCCGAGGAGGAGCCTGGAGCAGATTCTGACGCGAG 541
 QY 187 TyrThrThrPheValIysProAlaPheGluGluPheCysLeuProThrIlyIlyValIle 206
 DB 542 TACACACCTCTCTGTAAGCGCGCTTCGAGAGATTGCTGCCGCAAGAAAGATATGCC 601
 QY 207 AspValIleIleProArgGlyVal 214
 DB 602 GATGTGATCATCCCGCAGAGAGTT 625
 RESULT 12
 B1261258 827 bp mRNA linear EST 17-JUL-2001
 LOCUS B1261258
 DEFINITION 602969110.F1 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:5108666 5',
 mRNA sequence.
 ACCESSION B1261258
 VERSION B1261258.1 GI:14820349
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE NIH-MGC http://mgs.nci.nih.gov/.
 1 (bases 1 to 827)
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 TITLE Unpublished
 JOURNAL
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: Incyte Genomics, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: LHAM1263 row: c column: 03
 High quality sequence start: 5
 High quality sequence stop: 799.
 Location/Qualifiers
 1..827
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5108666"
 /tissue_type="cervical carcinoma cell line"
 /lab_host="DH10B"
 /clone_lib="NIH_MGC_12"
 /note="Organ: cervix; Vector: pCMV-SPORT6; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.

Average insert size 1.4 kb. Library prepared by Life Technologies "Technology" 189 a 210 c 274 g 154 t

Alignment Scores:

Pred. No.: 5.52e-195 Length: 827
Score: 205.00 Matches: 218
Percent Similarity: 99.54% Conservative: 0
Best Local Similarity: 99.54% Mismatches: 0
Query Match: 74.01% Indels: 1
DB: 12 Gaps: 0

US-09-896-522-2 (1-277) x B1261258 (1-827)

Qy 1 MetAlaSerAlaGlyGlyGluAAspCySerProAlaProGluAlaAspArgProHis 20
Db 52 ATGGCTTCGGCGGGGGAAGACTGCGAGAGCCCGCGCGGAGCGACCGTCCGCAC 111
Qy 21 GlnArgProPheLeuIleGlyValSerGlyThrAlaSerGlyLysSerThrValCys 40
Db 112 CAGCGGCCCTTCCTGATAGGGGTGAGCGCGGCACTGCCAGCGGAAGTCCAGCTGTGT 171
Qy 41 GlnLysIleMetGluLeuLeuGlyGlnAsnGluValGluGlnArgGlnArgLysVal 60
Db 172 GAGAGATCATGAGATTCTCGGACAGAACAGAGGTGAAACGCGGACGCGAAGGTGTC 231
Qy 61 IleLeuSerGlnAspArgPheTyrlsValLeuThrAlaGluGlnLysAlaLysAlaLeu 80
Db 232 ATCCGTGACCAGAGAGGTTCTACAAAGTCTCTGACGCGAGAGAGCCAGGCCCTTG 291
Qy 81 LysGlyGlnTyraAspPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThr 100
Db 292 AAAGGACAGTCAATTTTGACCATCCAGATGCCCTTGATATGATTGATCACAAGACT 351
Qy 101 LeuLysAsnIleValGluGlyLysThrValGluValProThrTyraAspPheValThrHis 120
Db 352 CTGAAGAACATCGTGAAGGCAAAACGGGTGAGGCGGACCTATGATTTGTGACACAC 411
Qy 121 SerArgLeuProGluThrThrValValTyrlProAlaAspValValLeuPheGluGlyIle 140
Db 412 TCAAGGTTACAGAGACACAGGTGTCTACCCCTGCGGACGCGGTCTGTTGAGGGATC 471
Qy 141 LeuValPheTyrlSerGlnGluIleArgAspMetPheHisLeuArgLeuPheValAspThr 160
Db 472 TTGGTTGTTCTACACCGAGAGATCCGGACATGTTCCACCTGCCCTCTTGTGACAC 531
Qy 161 AspSerAspValArgLeuSerArgArgValLeu-ArgAspValArgArgGlyArgAspLe 180
Db 532 GACTCCGACGTCAAGCTGTCTCGAAAGATTCTCCCGGAGCGTGGGAGGAGGACT 591
Qy 180 uGluGlnIleLeuThrGlnTyrlThrThrPheValLysProAlaPheGluGluPheCysLe 200
Db 592 GGAACAGATTTGACGAGTACACCACTTGTGAAGCGGCGCTTCGAGGGTTCGTGCT 651
Qy 200 uProThrLysLysTyrlAlaAspValIleIleProArgGlyValAspAsnMetVal 218
Db 652 GCCGACAAAGAGATATGCCGATGTGATCATCCACGAGAGTGGACATATAGTT 706

RESULT 13
BG491358 860 bp mRNA linear EST 27-MAR-2001
LOCUS 60253642p1 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:4684738 5',
DEFINITION mRNA sequence.
ACCESSION BG491358
VERSION BG491358.1 GI:13452870
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 860)
AUTHORS NIH-MGC <http://mhc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-rc@mail.nih.gov
Tissue Procurement: DCTD/DRP
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: LLCM493 row: c column: 11
High quality sequence stop: 844.
Location/Qualifiers

FEATURES

source

1..860
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4684738"
/tissue_type="amelanotic melanoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_id="NIH_MGC_41"
/note="Organ: skin; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; CDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
Laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH MGC library."

BASE COUNT

195 a 223 c 288 g 154 t

ORIGIN

Alignment Scores:
Pred. No.: 5.83e-193 Length: 860
Score: 203.00 Matches: 216
Percent Similarity: 99.54% Conservative: 0
Best Local Similarity: 99.54% Mismatches: 0
Query Match: 73.29% Indels: 1
DB: 10 Gaps: 0

US-09-896-522-2 (1-277) x BG491358 (1-860)

Qy 1 MetAlaSerAlaGlyGlyGluAAspCySerProAlaProGluAlaAspArgProHis 20
Db 52 ATGGCTTCGGCGGGGGAAGACTGCGAGAGCCCGCGGAGCGACCGTCCGCAC 111
Qy 21 GlnArgProPheLeuIleGlyValSerGlyThrAlaSerGlyLysSerThrValCys 40
Db 112 CAGCGGCCCTTCCTGATAGGGGTGAGCGCGGCACTGCCAGCGGAAGTCCAGCTGTGT 171
Qy 41 GlnLysIleMetGluLeuLeuGlyGlnAsnGluValGluGlnArgGlnArgLysVal 60
Db 172 GAGAGATCATGAGATTCTCGGACAGAACAGAGGTGAAACGCGGACGCGAAGGTGTC 231
Qy 61 IleLeuSerGlnAspArgPheTyrlsValLeuThrAlaGluGlnLysAlaLysAlaLeu 80
Db 232 ATCCGTGACCAGAGAGGTTCTACAAAGTCTCTACCGCAGAGAGAGCCAGGCCCTTG 291
Qy 81 LysGlyGlnTyraAspPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThr 100
Db 292 AAAGGACAGTCAATTTTGACCATCCAGATGCCCTTGATATGATTTGATGACAGACT 351
Qy 101 LeuLysAsnIleValGluGlyLysThrValGluValProThrTyraAspPheValThrHis 120
Db 352 CTGAAGAACATCGTGAAGGCAAAACGGGTGAGGCGGACCTATGATTTGTGACACAC 411
Qy 121 SerArgLeuProGluThrThrValValTyrlProAlaAspValValLeuPheGluGlyIle 140
Db 412 TCAAGGTTACAGAGACACAGGTGTCTTACCTCGGAGAGTGTCTTGTGAGGGATC 471
Qy 141 LeuValPheTyrlSerGlnGluIleArgAspMetPheHisLeuArgLeuPheValAspThr 160

[illegible]

/clone_11b="NIH_MGC_12"
/note="Organ: cervix; Vector: PCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.4 kb. Library prepared by Life
Technologies."

BASE COUNT 196 a 219 c 271 g 162 t
ORIGIN

Alignment Scores:

Pred. No.:	5.98e-187	Length:	848
Score:	197.00	Matches:	210
Percent Similarity:	99.53%	Conservative:	0
Best Local Similarity:	99.53%	Mismatches:	0
Query Match:	71.12%	Indels:	1
DB:	12	Gaps:	0

US-09-896-522-2 (1-277) x BI258532 (1-848)

```
QY 37 SerThrValCysGluLysIleMetGluLeuGluGlnAsnGluValGluGlnArgGln 56
   |||||
Db 87 TCGACCGGTGTGAGAGATCATGGAGTTGCTGGGACAGACGAGTGGACACGGCGCAG 146
QY 57 ArgLysValValIleLeuSerGlnAspArgPheTyrLysValLeuThrAlaGluGlnLys 76
   |||||
Db 147 CGGAGGTGTGATCTGTGAGCAGGACAGGAGTTCTTACAGGTCCTGACGCGCAGAGAG 206
QY 77 AlaLysAlaLeuLysGluGlnTyrAsnPheAspHisProAspAlaPheAspAsnAspLeu 96
   |||||
Db 207 GCCAAGGCTTGAAGAGACATGACATTTTGAACATCCAGATGCCATTGATATGATTTG 266
QY 97 MethIAsnGlnLeuLysAsnIleValGluGlnLysThrValGluValProThrTyrAsp 116
   |||||
Db 267 ATGCACAGACTCTGAAGAACATCGTGGAGGGCAAAACGGTGGAGGTGCCGACCTATGAT 326
QY 117 PheValThrHisSerArgLeuProGluThrThrValValTyrProAlaAspValValLeu 136
   |||||
Db 327 TTTGTGACACCTCAAGGTACCAAGAGCCACGGTGTCTACCTCGCGAGCGTGTCTG 386
QY 137 PheGluGlyIleLeuValPheTyrSerGlnGluIleArgAspMetPheHisLeuArgLeu 156
   |||||
Db 387 TTTGAGGGCATCTTGGTGTCTTACAGCCAGAGATCCGGACATGTTCCACCTGCGCTC 446
QY 157 PheValAspThrAspSerAspValArgLeuSerArgValLeuArgAspValArgArg 176
   |||||
Db 447 TTCGTGACACCGACTCGAGCTGTCTCGAAGAGTTCTCCGGAGCTGCGCGA 506
QY 177 GluArgAspLeuGluGlnIleLeuThrGlnTyrThrPheValLysProAlaPheGlu 196
   |||||
Db 507 GGGAGGAGCTCGAGCATTTCTGACGACATCACACCTTCGTGAAGCCGGCCTTCGAG 566
QY 197 GluPheCysLeuProThrLysLysTyrAlaAspValIleIleProArgGlyValAspAsn 216
   |||||
Db 567 GAGTTCGTGCTCGCGACCAAGACATATGCCATGTGATCATCCACAGAGATGACAT 626
QY 217 MetValAlaIleAsnLeuIleValGlnHisIleGlnAspIleLeuAsn-GlyAspIleCys 236
   |||||
Db 627 ATGGTGCATCAACATGATCGTGCACACATCCAGGACATTTGAAATGGTGTGACATCTG 686
QY 236 GlyTyrHisArgGlyGlySerAsnGlyArg 246
   |||||
Db 687 CAAATGGACCGAGAGGATCAATGGGCGG 717
```

Search completed: November 25, 2003, 12:09:32
Job time : 2555 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 25, 2003, 10:20:53 ; Search time 74 Seconds
(without alignments)
1652.204 Million cell updates/sec

Title: US-09-896-522-2
Perfect score: 277
Sequence: 1 MASAGDECSPPAPADRP.....HPGMLTSGKSHLESSSRPH 277

Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 56978 seqs, 220691566 residues

Word size: 15

Total number of hits satisfying chosen parameters: 2

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODE=frame+;p2n.model -DEV=x1h
-Q=/cgn2_1/USPTO.epool/US09895522/rnucat_21112003_184500_5896/app_query.fasta.1.455
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=p2n_ol1.rn1 -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=qual:cy -THR MIN=15 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09895522.QCGN_1_1_56.0rnucat_21112003_184500_5896 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEOQUERY -NEG SCORES=0 -WAIT -DSBLOK=100 -LONGLOG
-DEV TIMBOU=120 -WARN TIMBOU=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA.*
1: /cgn2_6/prodata/2/ina/5A_COMB.seq.*
2: /cgn2_6/prodata/2/ina/5B_COMB.seq.*
3: /cgn2_6/prodata/2/ina/5A_COMB.seq.*
4: /cgn2_6/prodata/2/ina/5B_COMB.seq.*
5: /cgn2_6/prodata/2/ina/PCTUS_COMB.seq.*
6: /cgn2_6/prodata/2/ina/backfillseq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	277	100.0	834	4	US-09-536-647-1
2	277	100.0	1288	4	US-09-620-312D-546

ALIGNMENTS

RESULT 1
US-09-536-647-1
; Sequence 1, Application US/09536647
; Patent No. 6579708
; GENERAL INFORMATION:
; APPLICANT: Ho, Yen Sen

; APPLICANT: Johnson, Randall
; TITLE OF INVENTION: No. 6579708e1 Human Uridine Kinase
; FILE REFERENCE: GP50020
; CURRENT APPLICATION NUMBER: US/09/536,647
; CURRENT FILING DATE: 2000-03-27
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 834
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(780)
US-09-536-647-1

Alignment Scores:
Pred. No.: 1.75e-279 Length: 834
Score: 277.00 Matches: 277
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0

US-09-896-522-2 (1-277) x US-09-536-647-1 (1-834)

QY	1	MetAlaserAlaGlyGlyGluAspCysGluSerProAlaProGluAlaAspArgProHis	20
DB	1	ATGGCTTCGGCGCGGAGCGCAACATCGCGAGCCCCCGCGGAGCCGACCGTCGCGCAC	60
QY	21	GlnArgProPheLeuIleGlyValSerGlyThrAlaSerGlyLysSerThrValCys	40
DB	61	CAGCGGCCCTTCCTGATAGAGGGGTAGCGCGGCACTCCAGCGGAGAGTCGACCGTGT	120
QY	41	GluValIleMetGluLeuLeuGlyGlnAsnGluValGluGlnArgGlnArgValVal	60
DB	121	GAGAAAGATCATGAGTGTGCGGACAGAAACGAGTGAACACGCGGACGCGGAGGTC	180
QY	61	IleLeuSerGlnAspArgPheTyrIleValIleThrAlaGluGlnLysAlaLysAlaLeu	80
DB	181	ATCCCGAGCGCAGACAGGTTCTACAAAGTCTCGACGCGCAGACCAAGGCGCTTG	240
QY	81	LysGlyGlnTyrAsnPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThr	100
DB	241	AAAGGACGTACATTTTGAACCATCAAGATGCTTTGACAAATGATTTGATGACAGACT	300
QY	101	LeuLysAsnIleValGluGlyLysThrValGluValProThrTyrAspPheValThrHis	120
DB	301	CTGAAGAACATGCTGAGAGGCAAAACGGTGGAGGTCGACCTATGATTTGTGACACAC	360
QY	121	SerArgLeuProGluThrThrValValTyrProAlaAspValValLeuPheGluGlyIle	140
DB	361	TCAAGTTTACAGAACACCGGTGCTTACCTCGCGGACGTTGTTGTGAGGGGCATC	420
QY	141	LeuValPheTyrSerGluGluIleLeuArgMetPheHisLeuArgLeuPheValAspThr	160
DB	421	TGGGTGTTTACAGCCAGAGATCCGGACATGTTCCACTGCGCTCTTCTGTGACACC	480
QY	161	AspSerAspValArgLeuSerArgArgValIleuArgAspValArgArgGlyArgAspLeu	180
DB	481	GACTCGACGTACAGGCTGTCTGAAAGATTCCTCCGGAGCGTCCGAGGAGGAGACCTG	540
QY	181	GluGlnIleLeuThrGlnTyrThrThrPheValLysProAlaPheGluGluPheCysLeu	200
DB	541	GAGCAGATTCTGACCCAGACACACCTTCGTAAGACCGGCTTGAGAGGATTCGCTCG	600
QY	201	ProThrLysTyrAlaAspValIleIleProArgGlyValIaAspAsnMetValAlaIle	220
DB	601	CCGACAAAGAAAGTATGCGGATGTATCATCCACAGAGAGTGACAAATATGTTGCATC	660
QY	221	AsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysLysTyrHisArg	240
DB	661	AACCTGATGTCGACACATCCAGACATTTGGAATGTGATCATCTGCAAAATGGACCGCA	720

```

QY      241  G|YGIYSerAenGIYArgSerTYrLYsArgThrPheSerGIuProGIYAspHisProGIY 260
Db      721  GGAAGGTCCTCAATGGCGAGCTACAGCGACCTTTCTTGAGCCAGGAGACACCTGGG 780

QY      261  MetLeuThrSerGIYLYsArgSerHisLeuGIuSerSerSerAArgProHis 277
Db      781  ATGCTGACCTCTGGCAAAAGGTCACATTGGAGTCCAGACAGACACCCAC 831

RESULT 2
US-09-620-312D-546
; Sequence 546, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yungqing
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6569662e1 Nucleic Acids and
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: PL_FL_genes Version 1.0
; SEQ ID NO 546_
; LENGTH: 1288 .
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (95)..(928)
US-09-620-312D-546

Alignment Scores:
Pred. No.:      2,64e-279      Length:      1288
Score:          277.00         Matches:      277
Percent Similarity: 100.00%     Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match:     100.00%       Indels:      0
DB:              4            Gaps:        0

US-09-896-522-2 (1-277) x US-09-620-312D-546 (1-1288)

QY      1  MetAlaSerAlaGIYGIYAspCYsGIuSerProAlaProGIYAlaAspArgProHis 20
Db      95  ATGGCTTGGGGGAGGGAAGACTGAGAGCCCGCGCGAGCGACCGTCCGCAC 154

QY      21  GlnArgProPheLeuIleGIYValSerGIYGIYThrAlaSerGIYLYsSerThrValCys 40
Db      155  CAGCGGCGCTTCCGTGATAGGGGTGAGCGCGCGACCTGCCAGCGGAAAGTGCACCGTGTGT 214

QY      41  GluLYsIleMetGIuLeuLeuGIYGIYAsnGIYValGIuGIuArgGIYAlaArgLYsValVal 60
Db      215  GAGAAAGATCATGAGATTGCTGTGGACAGAAAGAGGTGAAACAGCGGACGGAAGTGTGTC 274

```

```

QY      61  IleLeuSerGIuAspAArgPheTYrLYsValIleuThrAlaGIuGIuLYsAlaLYsAlaLeu 80
Db      275  ATCTGAGCCAGGACAGGTTCTTACAGTCTTACAGGACGAGACAAAGGCCAGGCTTG 334

QY      81  LYsGIYGIYLYsArgPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThr 100
Db      335  AAGGACAGTACATTTTGGATCCATCCAGATGCTTTGATATATGATTGATGACACAGACT 394

QY      101  LeuLYsAsnIleValGIuGIYLYsThrValGIuValProThrTYrAspPheValThrHis 120
Db      395  CTGAAAGACATCTGTGAGGCGCAAAAGGTGAGAGTCCACCTATATTTTGTGACACAC 454

QY      121  SerArgLeuProGIuThrThrValValTYrProAlaAspValIleuPheGIuGIYIle 140
Db      455  TCAGGTTACAGAGACACAGGTGTGTCTTACCTCGCGAGCTGTTCTTTGAGGACATC 514

QY      141  LeuValPheTYrSerGIuGIYLYsArgAspMetPheHisLeuArgLeuPheValAspThr 160
Db      515  TTGTTGTTCTACAGCGAGGATCCGGACATGTTCACCTGGCGCTTCTCGTGACACC 574

QY      161  AspSerAspValArgLeuSerArgValLeuArgAspValArgArgGIYArgAspLeu 180
Db      575  GACTCCGACGTCAAGCTGTCTCGAAGAGTCTCCGGAGGTGCGCCGAGGAGGACCTG 634

QY      181  GluGIuIleLeuThrGIYTYrThrThrPheValLYsProAlaPheGIuGIuPheCysLeu 200
Db      635  GACGAGTTCTGACGACGATCACACCTTCGTGAAAGCCGGCTTCCAGAGATTTCTCCTG 694

QY      201  ProThrLYsLYsTYrAlaAspValIleIleProArgGIYValAspAsnMetValAlaIle 220
Db      695  CCGACAAAGAAATGATCCGATGTGATCATCCACGAGAGTGGACAATATGTTGTCATC 754

QY      221  AsnLeuIleValGIuIleIleGlnAspIleLeuAsnGIYAspIleCysLYsTYrPheArg 240
Db      755  AACCTGATCGTGACACATCCAGACATTTGAAATGTGACATTCGCAATGGCACCGA 814

QY      241  G|YGIYSerAenGIYArgSerTYrLYsArgThrPheSerGIuProGIYAspHisProGIY 260
Db      815  GAGGGTCCAAATGGGCGGACCTTCAAGCGGACCTTTCTGAGCCAGAGGAGACCCCTGGG 874

QY      261  MetLeuThrSerGIYLYsArgSerHisLeuGIuSerSerSerAArgProHis 277
Db      875  ATGCTGACCTCTGGCAAAAGGTCACATTGGAGTCCAGACAGACCCAC 925

```

Search completed: November 25, 2003, 12:10:48
Job time : 76 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: November 25, 2003, 08:18:06 ; Search time 318 Seconds

(without alignments)
2351.397 Million cell updates/sec

Title: US-09-896-522-2

Perfect score: 277

Sequence: 1 MASAGDEDESPAPFADRP.....HPGMLTSGKSHLESSSRPH 277

Scoring table: OLIGO

Xgapop 60.0 , Xgapext 60.0

Ygapop 60.0 , Ygapext 60.0

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 2552756 segs, 1349719017 residues

Word size: 15

Total number of hits satisfying chosen parameters: 30

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: listing first 45 summaries

Command line parameters:

-MODE=frame+.p2n.model -DEV=xlh

-Q=/c9n2.1/USPTO/US09896522/rnatc.21112003.184458.5860/app_query.fae.1.455

-DB=N.Geneseq.19Jun03 -QPMT=fstap -SUFIX=p2n.oli.rng -MINMATCH=0.1 -LOOPCL=0

-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human0.cdi

-LIST=45 -DOCALLIGN=200 -THR SCORE=quality -THR MIN=15 -ALIGN=15 -MODE=LOCL

-OUTPM=pro -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000

-USER=US09896522.@CGN.1.1.312.@runat.21112003.184458.5860 -NCPU=6 -ICPU=3

-NO_MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGALIGN

-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREDS=1 -XGAPOP=60 -XGAPEXT=60 -FCGAPOP=6

-FCGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

1: N.Geneseq.19Jun03.*

2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*

3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*

4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*

5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*

6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*

7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*

8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*

9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*

10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*

11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*

12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*

13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*

14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*

15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*

16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*

17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*

18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*

19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*

20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*

21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*

22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*

23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*

24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

25: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.*

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	277	100.0	834	23	AAK98735	DNA of a human uri
2	277	100.0	1288	22	AA158658	Human polynucleoti
3	277	100.0	1624	22	AA158658	Human 57658 DNA.
4	277	100.0	2152	22	AA160444	Human polynucleoti
5	277	100.0	2160	22	AA115621	Human cDNA sequenc
6	191	69.0	753	22	AA104832	Human cDNA clone (
7	189	68.2	900	22	AA175355	Human uridine kina
8	136	49.1	1396	22	AA189762	Human polynucleoti
9	112	40.4	734	24	AA151801	Human mdtl cDNA in
10	60	21.7	9732	22	AA136439	Human musculoskele
11	60	21.7	9732	22	ABX59427	CDNA encoding nove
12	60	21.7	19125	22	AA136440	Human musculoskele
13	60	21.7	19125	25	ABX59428	CDNA encoding nove
14	25	9.0	715	20	AA15454	Human gene express
15	25	9.0	1310	21	AA151785	Human prostate can
16	25	9.0	1322	22	AA123801	Human transferrase
17	25	9.0	1322	22	AB054470	Human ovarian anti
18	25	9.0	1402	22	AA157850	Human polynucleoti
19	25	9.0	3881	23	ABV23998	Human prostate exp
20	25	9.0	3881	23	ABV24102	Human prostate exp
21	25	9.0	3881	23	ABV24281	Human prostate exp
22	25	9.0	3881	23	ABV29881	Human prostate exp
23	25	9.0	3881	23	ABV29983	Human prostate exp
24	25	9.0	3881	23	ABV30091	Human prostate exp
25	15	5.4	65	24	ABN27650	Rat spliced trans
26	15	5.4	65	24	ABN51411	Mouse spliced tran
27	15	5.4	549	23	ABT33606	Anticancer agent t
28	15	5.4	822	23	AB110981	Drosophila melanog
29	15	5.4	1116	22	AA159636	Human polynucleoti
30	15	5.4	3413	23	AB110980	Drosophila melanog

ALIGNMENTS

RESULT 1	AAK98735	standard; DNA; 834 BP.
ID	AAK98735	
XX	AAK98735;	
AC		
XX		
DT	02-MAY-2002 (first entry)	
XX		
DE	DNA of a human uridine kinase (UDK).	
XX		
KW	Human; uridine kinase; diagnostic assay; mutation detection; UDK;	
KW	probe; chromosome localisation study; tissue expression; gene therapy;	
KW	antibody; vaccine; human ovarian cancer; immunological disorder;	
KW	human colon carcinoma; immunogen; ds.	
XX		
OS	Homo sapiens.	
XX		
FH	Key	Location/Qualifiers
FT	CDS	1..780
FT		/tag= a
FT		/partial
FT		/note="No stop codon"
XX		
PN	WO200172963-A2.	
XX		
PD	04-OCT-2001.	
XX		
PF	27-MAR-2001; 2001WO-US09663.	
XX		
PR	27-MAR-2000; 2000US-0536647.	
XX		
PA	(SMIK) SMITHKLINE BEECHAM CORP.	

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed,

XX Ho YS, Johnson RK;
 PI WPI; 2001-626259/72.
 XX P-PSDB; AAO14412.
 XX Novel human uridine kinase polypeptides useful for treating cancers,
 PT and to identify agonists and antagonists of the polypeptide useful for
 PT treating conditions associated with uridine kinase imbalance
 XX Claim 5; Page 22-23; 31pp; English.
 XX The invention relates to newly identified human uridine kinase (UDK)
 CC polypeptides and polynucleotides and methods for producing such
 CC polypeptides by recombinant techniques. Also disclosed in the invention
 CC are methods for utilizing uridine kinase polypeptides and polynucleotides
 CC in diagnostic assays. The polynucleotides and polypeptides of the
 CC invention may be used as diagnostic reagents by detecting mutations in an
 CC associated gene. An array of oligonucleotide probes comprising the
 CC uridine kinase polynucleotide sequence or fragments thereof can be
 CC constructed to conduct efficient screening of genetic mutations, for
 CC example. Detection of abnormally decreased or increased levels of
 CC polypeptide or mRNA expression may also be used for diagnosing or
 CC determining susceptibility of a subject to a disease of the invention.
 CC The polynucleotide sequences of the invention can be used for chromosome
 CC localisation studies and tissue expression studies. The polypeptides of
 CC the invention or fragments thereof may be used as immunogens to produce
 CC antibodies. These antibodies may be employed to isolate or identify
 CC clones expressing the polypeptide. The polypeptides and polynucleotides
 CC of the invention can be used as a vaccine or in gene therapy to treat
 CC diseases such as human ovarian cancer, human colon carcinomas, and
 CC immunological disorders. This polynucleotide sequence represents the DNA
 CC of a human uridine kinase of the invention.
 XX Sequence 834 BP; 194 A; 219 C; 265 G; 156 T; 0 other;
 SQ
 Alignment Scores:
 Pred. No.: 2e-269 Length: 834
 Score: 277.00 Matches: 277
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: Gaps: 0
 US-09-896-522-2 (1-277) x AAK98735 (1-834)
 QY 1 MetAlaSerAlaGlyGluAspCysGluSerProAlaProGluAlaAspArgProHis 20
 Db 1 ATGCTTCGGGCGGAGGAAAGACTGGGAGAGCCCGCGGAGCGCACCGTCCGCAC 60
 QY 21 GlnArgProPheLeuIleGlyValSerGlyThrAlaSerGlyLysSerThrValCys 40
 Db 61 CAGGGCGCCCTTCCGATAGGGGTGAGCGCGGCACCTGCCAGCGGAAGTCGACCTGTGT 120
 QY 41 GluLysIleMetGluLeuLeuGlyGlnAsnGluValGluGlnArgGlnAsnGlyValVal 60
 Db 121 GAGAAAGATCATGAGTCTGCGACAGAACAGAGGTGGACAGCGCGAGCGAAGTGTGTC 180
 QY 61 IleLeuSerGlnAspArgPheTyrLysValLeuThrAlaGluGlnLysAlaLysValLeu 80
 Db 181 ATCTGAGCCAGCAGAGGTTCTTACAAGTCTCTGACGGCAGAGCAAGAGCGCAAGCCTTG 240
 QY 81 LysGlyGlnTyrAsnProPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThr 100
 Db 241 AAAAGACAGTCAATTTTGAACATCCAGATGCTTTTACATGATTTGATGACACAGACT 300
 QY 101 LeuLysAsnIleValIleGluGlyLysThrValGluValProThrTyrAspPheValHis 120
 Db 301 CTGAAAGACATCGAGAGGCAAAACGGTGGAGGTGCCGACCTTATGATTTTGTACACAC 360
 QY 121 SerArgLeuProGluThrThrValValTyrProAlaAspValValLeuPheGluGlyIle 140
 Db 361 TCAAGGTTTACAGAGACCAAGCGGTGTCTACCTCTCGGACGTGTCTTGTGAGGCGATC 420

QY 141 LeuValPheTyrSerGlnGluIleArgAspMetPheHisLeuArgLeuPheValAspThr 160
 Db 421 TTGGTGTCTTACAGCCAGAGATCCGGGACATGTTCCACTGCGCTTCTGGACACC 480
 QY 161 AspSerAspValArgLeuSerArgValLeuArgAspValArgArgGlyLysArgAspLeu 180
 Db 481 GATTCGACGCTCAGGCTGTCTCGAAGAGTTCTCCGGACGTGCGCGAGAGGAGACCTG 540
 QY 181 GluGlnIleLeuThrGlnTyrThrThrPheValLysProAlaPheGluGluPheCysLeu 200
 Db 541 GACGAGTTCTGACGCGATACACCACTTGTAAGCCGGCCTTCAGAGAGTTCTGCTG 600
 QY 201 ProThrLysLysTyrThrAlaAspValIleIleProArgGlyValAspAsnMetValAlaIle 220
 Db 601 CCACAAAGAAAGATATCCGATGTATCATCCACGAGAGTGACAAATATGTGTGCATC 660
 QY 221 AsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysLysTrpHisArg 240
 Db 661 AACCTATGTCGACGACATCCAGACATTTCTGATGTGACATTCGAAATGGCACCGA 720
 QY 241 GlyGlySerAsnGlyArgSerTyrLysArgThrPheSerGluProGlyAspHisProGly 260
 Db 721 GAGGGTCCAAATGGCGGAGCTACAGCGGACCTTTTCTGAGCCAGGGAGCACCTGGG 780
 QY 261 MetLeuThrSerGlyLysArgSerHisLeuGluSerSerArgProHis 277
 Db 781 ATGCTGACCTCTGGCAAAAGGTCACTTGTGAGTCCAGACAGACCCAC 831
 RESULT 2
 AA158658
 ID AA158658 standard; cDNA, 1288 BP.
 AC AA158658;
 XX 22-OCT-2001 (first entry)
 DT
 XX
 DE Human polynucleotide SEQ ID NO 861.
 KW Human; nocotropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 KW leukemia; ss.
 KW
 XX
 OS Homo sapiens.
 XX
 PN WO200153312-A1.
 XX
 PD 26-JUL-2001.
 XX
 XX 26-DEC-2000; 2000WO-US34263.
 XX
 PR 21-JAN-2000; 2000US-0488725.
 PR 25-APR-2000; 2000US-0552317.
 PR 09-JUL-2000; 2000US-0598042.
 PR 19-JUL-2000; 2000US-0620312.
 PR 03-AUG-2000; 2000US-0653450.
 PR 14-SEP-2000; 2000US-0662191.
 PR 19-OCT-2000; 2000US-0693036.
 PR 29-NOV-2000; 2000US-0727344.
 XX
 XX (HYSEQ-) HYSEQ INC.
 PA
 XX
 PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
 XX WPI; 2001-442253/47.
 DR P-PSDB; AAM39502.
 XX
 PT Novel nucleic acids and polypeptides, useful for treating disorders

PT such as central nervous system injuries -
 XX
 PS Claim 1; SEQ ID NO 861; 10078bp; English.
 XX
 CC The invention relates to human nucleic acids (AA157798-AA161369) and
 CC the encoded polypeptides (AAW38642-AAW42213) with nootropic,
 CC immunosuppressant and cytoprotective activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,
 CC Actin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemia and
 CC C.N.S disorders.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification.
 CC
 XX
 SQ Sequence 1288 BP; 275 A; 361 C; 418 G; 234 T; 0 other;

Alignment Scores:

Pred. No.: 3e-269 Length: 1288
 Score: 277.00 Matches: 277
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 22 Gaps: 0

US-09-896-522-2 (1-277) x AA158658 (1-1288)

QY 1 MetAlSerAlaGlyGluAspCysGluSerProAlaProGluAlaAspArgProHis 20
 DB 95 ATGGCTTCGGCGGAGGCGGAAAGACTCGAGAGCCCGCGGAGGCGCGCTCCGAC 154
 QY 21 GlnArgProPheLeuIleGlyValSerGlyGlyThrAlaSerGlyLysSerThrValCys 40
 DB 155 CAGCGGCCCTTCCTGATAGGGGTGAGGGGGGAGCTGCGCGGAGAACTGACCGTGTG 214
 QY 41 GlnLysIleMetGlnLeuLeuGlyGlnAsnGluValGlnLysGlnArgLysValAla 60
 DB 215 GAGAAATATCATGAGTTGCTCGGACAGAAAGAGTGGAAACGGCGGACCGGAGGTGTC 274
 QY 61 IleLeuSerGlnAspArgPheTyrLysValLeuThrAlaGluGlnLysAlaLysAlaLeu 80
 DB 275 ATCTGAGCGCAGACAGGTTCTCAAGGTCCTGACGCGCAGACAGAAAGCCAGCCCTTG 334
 QY 81 LysGlyGlnTyrTrpAsnPheAspHisProAspAlaPheAspAsnLeuMetHisArgThr 100
 DB 335 AAAGGACAGTACAAATTTTGAACATCCAGATGCTTTGATATATATTGATGACAGGACT 394
 QY 101 LeuLysAsnIleValGlnGlyLysThrValGluValProThrTyrAspPheValThrHis 120
 DB 395 CTGAAAGAATCATCTGGAGGCGCAAAACGGTGAAGTCCGACCTATGATTTTGGACACAC 454
 QY 121 SerArgLeuProGluThrThrValValTyrProAlaAspValValLeuPheGlnGlyLe 140
 DB 455 TCNAAGGTTACCAAGACCAACGAGTGTCTACCTCGGAGCGTGTCTGTTGAGGCGATC 514
 QY 141 LeuValPheTyrSerGlnGlnIleArgAspMetPheHisIleuArgLeuPheValAspThr 160
 DB 515 TTGGTGTTCACACCAAGAGATCCGAGCATTTCCACCTGTGCTCTTCGTTGAGCACCC 574
 QY 161 AspSerAspValArgLeuSerArgArgValLeuArgAspValArgArgGlyArgAspLeu 180
 DB 575 GATTCGAGCTCAGGCTGTCTCGAAGAGTTCTCCGGGAGCTGCGCGAGGAGGACCTG 634
 QY 181 GlnGlnIleLeuThrGlnTyrThrThrPheValLysProAlaPheGlnGluPheCysLeu 200
 DB 635 GAGCAGATTCTGACGAGTACACACACCTTCGTGAACCGGCTTCGAGGAGTCTGCTG 694

QY 201 ProThrLysLysTyrAlaAspValIleIleProArgGlyValAspAsnMetValAlaIle 220
 DB 695 CCGACAAAGAAAGTATGCCATGTGATCATCCACAGAGGTGAGCAATATGTGTCATC 754
 QY 221 AsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysLysThrHisArg 240
 DB 755 AACCTGATGTCGACGACATCCAGACGACATTCGAATGTGATGATCTGCAATGGCACCGA 814
 QY 241 GlyGlySerAsnGlyArgSerTyrLysArgThrPheSerGlnProGlyLysPheProGly 260
 DB 815 GAGGGGTCCAAATGGGGGAGCTTCAACAGCGACCTTTTCTGAGCCAGGGGAGCACCCCTGG 874
 QY 261 MetLeuThrSerGlyLysArgSerHisLeuGluSerSerSerArgProHis 277
 DB 875 ATGCTGACCTCTGGCAACGTCATCATTTTGGAGTCCAGACGACACCCAC 925
 RESULT 3
 AAD27186
 ID AAD27186 standard; DNA; 1624 BP.
 XX
 AC AAD27186;
 XX
 DT 18-APR-2002 (first entry)
 XX
 DE Human 57658 DNA.
 XX
 KW Human; uridine kinase-like protein; haematopoietic neoplastic disorder;
 KW 57658 protein; leukaemia; haematopoietic disorder; diabetes mellitus;
 KW rheumatoid arthritis; psoriasis; dermatitis; osteoporosis; rickets;
 KW sarcoma; myocardial infarction; hypertension; atherosclerosis;
 KW Alzheimer's disease; Parkinson's disease; Huntington's disease;
 KW gene therapy; chromosome mapping; tissue typing; dermatological;
 KW cytotactic; osteopathic; cardiac; neuroprotective; nootropic;
 KW anticonvulsant; ds.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 94..927
 FT /tag= a
 FT /product= "Human 57658 protein"
 FT misc_feature 94..924
 FT /tag= b
 FT /note= "This region is specifically claimed as
 SEQ ID NO: 3 in claim 1 of the specification"
 FT
 PN WO200202761-A2.
 XX
 PD 10-JAN-2002.
 XX
 PF 28-JUN-2001; 2001WO-US21063.
 XX
 PR 30-JUN-2000; 2000US-216503P.
 XX
 PA (MILL-) MILLENNIUM PHARM INC.
 XX
 PI Glucksmann MA;
 XX
 DR MPI; 2002-140091/18.
 DR P-PSDB; AAE16592.
 XX
 PT New isolated human uridine kinase family polypeptide 57658, useful for
 PT treating hematopoietic neoplastic disorders and disorders of neurons,
 PT heart and blood vessels -
 PT
 PS Claim 1; Fig 1a; 103bp; English.
 PS
 CC The patent discloses human uridine kinase-like polypeptides, designated
 CC 57658 and polynucleotides encoding such proteins. 57658 DNAs and proteins
 CC are useful for developing diagnostic and therapeutic agents for 57658-
 CC mediated or related disorders such as haematopoietic neoplastic disorders
 CC (e.g. leukaemia), haematopoietic disorders (e.g. psoriasis, dermatitis,
 CC diabetes mellitus, rheumatoid arthritis, disorders of bone (e.g. rickets,

CC osteoporosis), heart (e.g. sarcoma, myocardial infarction), blood vessels
 CC (e.g. hypertension, atherosclerosis) and neurons (e.g. Huntington's
 CC disease, Alzheimer's disease, Parkinson's disease). 57658 sequences or
 CC their antibodies are useful in screening assays, detection assays (e.g.
 CC forensic biology) and predictive medicine (e.g. diagnostic assays).
 CC prognostic assays, and monitoring clinical trials and pharmacogenomics).
 CC They are useful as reagents for diagnosing and treating 57658-mediated
 CC disorders. 57658 DNAs are useful in gene therapy, in chromosome mapping,
 CC to identify an individual from a minute biological sample (tissue typing)
 CC and to aid in forensic identification of the biological sample. The
 CC present sequence is a DNA encoding human 57658 protein.

XX
 XX Sequence 1624 BP; 354 A; 427 C; 506 G; 337 T; 0 other;

Alignment Scores:

Pred. No.:	3,72e-269	Length:	1624
Score:	277.00	Matches:	277
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	24	Gaps:	0

US-09-896-522-2 (1-277) x AAD27186 (1-1624)

QY 1 MetAlaSerAlaGlyGluAspCysGluSerProAlaProGluAlaAspArgProHis 20
 DB 94 ATGGCTTCGGCGGAGGAGAGCTGCGAGAGCCCGCGGAGCGGACCGTCCGAC 153
 QY 21 GlnArgProPheLeuIleGlyValSerGlyGlyThrAlaSerGlyLysSerThrValCys 40
 DB 154 CAGCGCCCTTCCTGATGAGGGGTGAGCGCGGACCTGCGAGGAGGAGTGCACCGTGTGT 213
 QY 41 GlnuylIleMetGluLeuLeuGlyGlnAsnGluValGluGlnArgGlnAsnGlyValVal 60
 DB 214 GAGAGATCATGAGGAGTGTGGGACAGAAACGAGGTGGAACCGCGACGCGAAGGTGCTC 273
 QY 61 IleuSerGlnAspArgPheTyrlLeuValLeuThrAlaGluGlnLysAlaValAlaLeu 80
 DB 274 ATCTGAGCCGAGGACGGTTTCAAGGTCTTGACGCGAGGACGAGAAAGGCCAGGCTTGG 333
 QY 81 LysGlyGlnTyrlAsnAspPheHisProAspAlaPheAspAspMetHisArgThr 100
 DB 334 AAAGAGATGATCAATTTGACCATTCAGATGCTTTGATATGATTTGATGACAGAGACT 393
 QY 101 LeuLysAsnIleValGluGlyLysThrValGluValProThrTyrlAspPheValThrHis 120
 DB 394 CTGAAGAACATCGTGGAGGCGAAACGCTGGAGGTGCGACCTATGATTTGTGACACAC 453
 QY 121 SerArgLeuProGluThrThrValValTyrlProAlaAspValValLeuPheGluGlyIle 140
 DB 454 TCAAGGTTTACAGAGACCGAGTGTCTACCTCGGACGAGTCTGTTTGAAGGACATC 513
 QY 141 LeuValPheTyrlSerGlnGluIleArgAspMetPheHisLeuArgLeuPheValAspThr 160
 DB 514 TTGGTGTTCCTACAGCCAGAGATCCGGGACATGTTCCACCTCGGCTTGTGTGACACC 573
 QY 161 AspSerAspValAlaGluSerArgArgValLeuArgAspValAlaGluGlyValArgAspLeu 180
 DB 574 GACTCCGACGTCAGGCTGTCTCGAAGGTTCTCCGGACGTCGCCAGAGGAGGAGACTG 633
 QY 181 GlnuIleLeuThrGlnTyrlThrThrPheValLysProAlaPheGluGluPheCysLeu 200
 DB 634 GAGAGATTCCTGAGCGCATACACCTTCTGTAAGCCGCGCTTCGAGAGTTTGTGCTG 693
 QY 201 ProThrLysLysTyrlAlaAspValIleLeuProArgGlyValAlaAspMetValAlaIle 220
 DB 694 CCGCAAAAGAGTATGCGATGATGATATCCACGAGGAGTGCAATATATGTTGCCATC 753
 QY 221 AsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysLysThrPheArg 240
 DB 754 AACCTGATCGGCGACATCCAGACATTTTGATAGGTGACATCTCCAAATGGCACCGA 813
 QY 241 GlyLysLeuAsnGlyArgSerTyrlLeuArgThrPheSerGluProGluAspHisProGly 260

DB 814 GAGAGTTCATATGGCGGAGCTACAGCGGACCTTTTCTGAGCGAGGAGACCACTGGG 873

QY 261 MetLeuThrSerGlyLysArgSerHisLeuGluSerSerArgProHis 277
 DB 874 ATGCTGACCTCTGGCAACGGTTCACATTTGAGTTCAGACAGACAGACCCGAC 924

RESULT 4

AA160444
 ID AA160444 standard; cDNA; 2152 BP.

XX AA160444;

DE 22-OCT-2001 (first entry)

XX Human polynucleotide SEQ ID NO 4433.

XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
 XX peripheral nervous system; neuropathy; central nervous system; CNS;
 XX Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 XX amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 XX chemokine; thrombolytic; drug screening; arthritis; inflammation;
 XX leukaemia; ss.

XX Homo sapiens.

XX NO200153312-A1.

XX 26-JUL-2001.

XX 26-DEC-2000; 2000MO-US34263.

XX 21-JAN-2000; 2000US-0488725.

XX 25-APR-2000; 2000US-0552317.

XX 09-JUL-2000; 2000US-0598042.

XX 19-JUL-2000; 2000US-0620312.

XX 03-AUG-2000; 2000US-0653450.

XX 14-SEP-2000; 2000US-0662191.

XX 19-OCT-2000; 2000US-0693036.

XX 29-NOV-2000; 2000US-0727344.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 XX Wang J, Wang Z, Wehman T, Xu C, Xue AJ, Yang Y, Zhang J;
 XX Zhao QA, Zhou P, Goodrich R, Drmanac RT;

XX WPI; 2001-442253/47.

XX P-PDB; AAM41288.

XX Novel nucleic acids and polypeptides, useful for treating disorders
 XX such as central nervous system injuries -

XX Claim 1; SEQ ID NO 4433; 10078bp; English.

XX The invention relates to human nucleic acids (AA157798-AA161369) and
 XX the encoded polypeptides (AAM38642-AAM42213) with nootropic,
 XX immunosuppressant and cytostatic activity. The polynucleotides are useful
 XX in gene therapy. A composition containing a polypeptide or polynucleotide
 XX of the invention may be used to treat diseases of the peripheral nervous
 XX system, such as peripheral nervous injuries, peripheral neuropathy and
 XX localised neuropathies and central nervous system diseases, such as
 XX Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 XX lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 XX utilisation of the activities such as: Immune system suppression,
 XX activation/inhibition activity, chemotactic/chemokinetic activity, haemostatic
 XX and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 XX assays for receptor activity, arthritis and inflammation, leukaemias and
 XX C.N.S disorders.

XX Note: The sequence data for this patent did not form part of the printed
 XX specification.

XX Sequence 2152 BP; 468 A; 581 C; 660 G; 442 T; 1 other;

Alignment Scores:

Pred. No.:	4,84e-269	Length:	2152
Score:	277.00	Matches:	277
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	22	Gaps:	0

US-09-896-522-2 (1-277) x AA160444 (1-2152)

```

QY 1 MetAlaSerAlaGlyGlyGluAspCysGluSerProAlaProGluAlaAspArgProHis 20
DB 59 ATGGCTTCGGCGGAGGCGAAGACTGCGAGAGCCCGCGCGAGGCGCGACCGTCCGCAC 118
QY 21 GluArgProPheLeuIleGlyValSer-GlyGlyThraLaseGlyLysSerThrValCys 40
DB 119 CAGCGGCCCTTCCTGATAGGGGTGAGCGGCGCACTGCCAGCGGGAAGTCGACCGTGTGT 178
QY 41 GluLysIleMetGluLeuLeuGlyGluAsnGluValGluGluArgGluNargLysValVal 60
DB 179 GAGAGATCATGAGATTGCTGGGACAGACGAGGTGAAACGCGGACGCGAGAGTGTCTC 238
QY 61 IleLeuSerGluAspArgPheTyTyLysValLeuThrAlaGluGluAlaLysAlaLeu 80
DB 239 ATCTGAGCCAGACGACGAGTCTCAAGAGTCTCTGACGCGACGAGGCGCAAGGCGCTTG 298
QY 81 LysGlyGluIleTrpAsnPheAspHisProAspAlaPheAspAsnLeuMetHisLysThr 100
DB 299 AAGGACAGATACATTTTGAACCATCCAGATGCCCTTATATATGATTTGATGACAGACT 358
QY 101 LeuLysAsnIleValGluGlyLysThrValGluValProThrTrpAspPheValThHis 120
DB 359 CTGAGAGACATCTGGGAGGCGAACCCTGAGAGTGCACCTTATGATTTTGACACAC 418
QY 121 SerArgLeuProGluIleThrValValTyProAlaAspValValLeuPheGluGlyIle 140
DB 419 TCAAGGTTACAGAGACACACGAGTGTCTACCTGCGGACGAGTCTGTTGAGGCGATC 478
QY 141 LeuValPheTyTySerGluGluIleArgAspMetPheHisLeuArgLeuPheValAspThr 160
DB 479 TTGGTGTCTTACAGCCAGAGATCCGGGACATGTTCCACCTGCCCTCTTCTGTGACACC 538
QY 161 AspSerAspValArgLeuSerArgArgValLeuArgAspValAlaArgArgGlyArgAspLeu 180
DB 539 GACTCCGACGTCAGGCTGTCTCGAAGAGTCTCCGGGACGTCGCCGAGGAGGAGGACTG 598
QY 181 GluGluIleLeuThrGluIleTyThrThrPheValLysProAlaPheGluGluPheCysLeu 200
DB 599 GAGAGATTCTGACGCGACGTCACACACTTCTGTGAAGCCGCGCTTCGAGGAGTCTGCTG 658
QY 201 ProThrLysLysTyTyAlaAspValIleIleProArgGlyValAlaAspMetValAlaIle 220
DB 659 CCCACCAAGAGATATGCCGATGATATCCCGGAGAGATGAGCAATATGTTGTCATC 718
QY 221 AsnLeuIleValGluHisIleGluAsnIleLeuAsnGlyLysIleCysLysThrHisArg 240
DB 719 AACCTGATCGTGACGACATCCAGACATCTGATGATGATCAATGCAATGACACCGA 778
QY 241 GlyIleSerAsnGlyArgSerTyTyLysArgThrPheSerGluProGlyAlaAspHisProGly 260
DB 779 GAGAGGTCGAATGGCGGAGCTACCAAGGACCTTTCTGAGGACGAGGAGCAACCTCGGG 838
QY 261 MetLeuThrSerGlyLysArgSerHisLeuGluLysSerSerArgProHis 277
DB 839 ATGCTGACCTCTGGCAACGATCATTGGAAGTCCAGACGACAGACCCGCAC 889

```

RESULT 5

AAH15621
ID AAH15621 standard; cDNA; 2160 BP.
AC AAH15621;
XX

DT 26-JUN-2001 (first entry)

DE Human cDNA sequence SEQ ID NO:13951.

KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

OS Homo sapiens.

XX EP1074617-A2.

XX 07-FEB-2001.

XX 28-JUL-2000; 2000EP-0116126.

XX 29-JUL-1999; 99JP-0248036.

XX 27-AUG-1999; 99JP-0300253.

XX 11-JAN-2000; 2000JP-0118776.

XX 02-MAY-2000; 2000JP-0183767.

XX 09-JUN-2000; 2000JP-0241899.

XX (HELI-) HELIX RES INST.

XX Ota T, Isogai T, Nishikawa T, Hayaashi K, Saito K, Yamamoto J;

XX Iishi S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX WPI; 2001-318749/34.

XX PT Primer sets for synthesizing polynucleotides, particularly the 5602

XX full-length cDNAs defined in the specification, and for the detection

XX of full-length cDNAs of the abnormality of the proteins encoded by the

XX full-length cDNAs -

XX PT

XX RS Claim 8; SEQ ID 13951; 2537pp + CD ROM; English.

XX XX

XX The present invention describes primer sets for synthesizing 5602

XX full-length cDNAs defined in the specification. Where a primer set

XX comprises: (a) an oligo-dT primer and an oligonucleotide complementary

XX to the complementary strand of a polynucleotide which comprises one of

XX the 5602 nucleotide sequences defined in the specification, where the

XX oligonucleotide comprises at least 15 nucleotides; or (b) a combination

XX of an oligonucleotide comprising a sequence complementary to the

XX complementary strand of a polynucleotide which comprises a 5'-end

XX sequence and an oligonucleotide comprising a sequence complementary to a

XX polynucleotide which comprises a 3'-end sequence, where the

XX oligonucleotide comprises at least 15 nucleotides and the combination of

XX the 5'-end sequence/3'-end sequence is selected from those defined in

XX the specification. The primer sets can be used in antisense therapy and

XX in gene therapy. The primers are useful for synthesizing polynucleotides,

XX particularly full-length cDNAs. The primers are also useful for the

XX detection and/or diagnosis of the abnormality of the proteins encoded by

XX the full-length cDNAs. The primers allow obtaining of the full-length

CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and

CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to

CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632

CC represent oligonucleotides, all of which are used in the exemplification

CC of the present invention.

CC XX

CC XX Sequence 2160 BP; 457 A; 591 C; 671 G; 441 T; 0 other;

CC SQ

CC Alignment Scores:

CC Pred. No.: 4.85e-269 Length: 2160

CC Score: 277.00 Matches: 277

CC Percent Similarity: 100.00% Conservative: 0

CC Best Local Similarity: 100.00% Mismatches: 0

CC Query Match: 100.00% Indels: 0

CC DB: 22 Gaps: 0

CC US-09-896-522-2 (1-277) x AAH15621 (1-2160)

QY 1 MetAlaSerAlaGlyGlyGluAspCysGluSerProAlaProGluAlaAspArgProHis 20

DB 95 ATGGCTTCGGCGGAGGCGAAGACTGCGAGAGCCCGCGAGGCGCGACCGTCCGCAC 154

```

QY 21 GlnArgProPheLeuIleGlyValSerGlyGlyThrAlaSerGlyLysSerThrValCys 40
Db 155 CAGCGGCCCTTCTGTATAGGGGTGAGCCGCCGCACTGCCACGGGAATCCAGCTGTGT 214
QY 41 GluLysIleMetGluLeuLeuGlyGlnAsnGluValGluGlnArgGlnArgLysValVal 60
Db 215 GAGAGAGATCATGAGATTCTGTGGACAGAACGAGGTGGAACGCGCAGCGGAGAGGTGTC 274
QY 61 IleLeuSerGlnAspArgPheTyrTyrValLeuThrAlaGluGluGlnValAlaAlaLeu 80
Db 275 ATCTGAGCCAGAGAGGTTCTTCAAGATCCTGACGGAGAGCAGAGGCGAAGGCTTG 334
QY 81 LysGlyGlnTyrAsnPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThr 100
Db 335 AAGAGACAGTACAATTTTGCATCCATCAGATGCTTTGATATGATTTGATCAGACGACT 394
QY 101 LeuLysAsnIleValGluGlyLysThrValGluValProThrTyrAspPheValThrHis 120
Db 395 CTGAAGAACATCGTGAAGGCAAAACGTTGAGGTGCGACCTATGATTTTGTGACACAC 454
QY 121 SerArgLeuProGluThrThrValValTyrProAlaAspValIleLeuPheGluGlyIle 140
Db 455 TCAAGTTTACCAAGAACCAAGGTGTCTACCTCGGAGGTGTCTGTGTGTGAGGGCATC 514
QY 141 LeuValPheTyrSerGlnGluIleArgAspMetPheHisLeuArgLeuPheValAspThr 160
Db 515 TTGGTGTCTACAGCAGAGATCCGGGACATGTTCCACCTGGCGCTCTTGTGTGACACC 574
QY 161 AspSerAspValAlaGluSerThrArgValIleuArgAspValAlaGArgGlyArgAspLeu 180
Db 575 GACTCCGACGTGAGGCTGTCTCGAAGAGTTCTCCGGGACGTGCGCGCAGGAGGAGACTCG 634
QY 181 GluGlnIleLeuThrGlnTyrThrThrPheValLysProAlaPheGluGluPheGlyLeu 200
Db 635 GAGCAGATTCGAGCCGCACTACCACTTCGTGAAGCCGGCCCTTGAGAGATCTCGCTG 694
QY 201 ProThrLysLysTyrAlaAspValIleLeuProArgGlyValAlaAspMetValAlaIle 220
Db 695 CCGCAAAAGAGTATGCCGATGTGATCATCCACGAGGAGTGGCAATATGTTGCCATC 754
QY 221 AsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysLysTyrHisArg 240
Db 755 AACCTGATCGTGCAGCAGCATCCAGACATTTGATGTGATCTGCATAATGACACCGA 814
QY 241 GlyLysSerAsnGlyArgSerTyrLysArgThrPheSerGluProGlyAspHisProGly 260
Db 815 GGAAGGTCATATGGCGGAGCTTACAGCGGACCTTTTCTGAGCCAGGGGAGACCCCTGG 874
QY 261 MetLeuThrSerGlyLysArgSerHisLeuGluSerSerSerArgProHis 277
Db 875 ATGCTGACCTCTGGCAACGATCATTTGGAGTCCAGACAGACCCCCAC 925

RESULT 6
AAH04832
ID AAH04832 standard; cDNA; 753 BP.
XX
XX AAH04832;
XX
XX 26-JUN-2001 (first entry)
XX
XX Human cDNA clone (5'-primer) SEQ ID NO:1667.
XX
XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
XX Homo sapiens.
XX
XX EP1074617-A2.
XX
XX 07-FEB-2001.
XX
XX 28-JUL-2000; 2000EP-0116126.
XX
XX 29-JUL-1999; 99JP-0248036.

```

```

PR 27-AUG-1999; 99JP-0300253.
PR 11-JAN-2000; 2000JP-0118776.
PR 02-MAY-2000; 2000JP-0183767.
PR 09-JUN-2000; 2000JP-0241899.
XX
XX (HELI-) HELIX RES INST.
PI Oca T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ichii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
DR WPI; 2001-318749/34.
XX
XX Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs -
XX
XX Claim 1; SEQ ID 1667; 2537pp + CD ROM; English.
XX
XX The present invention describes primer sets for synthesizing 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dt primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination
CC of the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialized methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95693 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.
XX
XX SQ Sequence 753 BP; 158 A; 196 C; 261 G; 135 T; 3 other;
XX
XX Alignment Scores:
XX Pred. No.: 8,21e-183 Length: 753
XX Score: 191.00 Matches: 191
XX Percent Similarity: 100.00% Conservative: 0
XX Best Local Similarity: 100.00% Mismatches: 0
XX Query Match: 68.95% Indels: 0
XX DB: 22 Gaps: 0

US-09-896-522-2 (1-277) x AAH04832 (1-753)
QY 1 MetAlaSerAlaGlyGlyGluAspCysGluSerProAlaProGluAlaAspArgProHis 20
Db 95 ATGGCTTCGGCGGAGGCGGAAGACTCGAGAGCCCGCCCGGAGGCCACGCTCCGCAC 154
QY 21 GlnArgProPheLeuIleGlyValSerGlyGlyThrAlaSerGlyLysSerThrValCys 40
Db 155 CAGCGGCCCTTCTGTATAGGGGTGAGCCGCCGCACTGCCACGGGAATCCAGCTGTGT 214
QY 41 GluLysIleMetGluLeuLeuGlyGlnAsnGluValGluGlnArgGlnArgLysValVal 60
Db 215 GAGAGAGATCATGAGATTCTGTGGACAGAACGAGGTGGAACGCGCAGCGGAGAGGTGTC 274
QY 61 IleLeuSerGlnAspArgPheTyrLysValLeuThrAlaGluGluGlnValAlaAlaLeu 80
Db 275 ATCTGAGCCAGAGAGGTTCTTCAAGATCCTGACGGAGAGCAGAGGCGAAGGCTTG 334
QY 81 LysGlyGlnTyrAsnPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThr 100
Db 335 AAGAGACAGTACAATTTTGCATCCATCAGATGCTTTGATATGATTTGATCAGACGACT 394

```


QY 101 LeuValAsnIleValGluGlyValThrValGluValProThrTyrAspPheValThrHis 120
 DB 395 CTGAAGAACATCGTGGAGGCAAAACGGTGGAGGCGACCTATGATTTGTACACAC 454
 QY 121 SerArgLeuProGluThrThrValValTyrProAlaAspValValLeuPheGluGlyLe 140
 DB 455 TCAAGGTACAGAGACACGGTGTCTACCCCTCGGACGTCGTTCGTTGAGGCGATC 514
 QY 141 LeuValPheTyrSerGluGluIleArgAspMetPheHisLeuArgLeuPheValAspThr 160
 DB 515 TTGGTGTCTACACCGACGAGATCCGGGACATGTTCCACCTCGCCCTTCGTGAGACCC 574
 QY 161 AspSerAspValArgLeuSerArgValLeuArgAspValArgArgGlyArgAspLeu 180
 DB 575 GACTCCACGTCAGCGTGTCTCGAAGAGTTCCTCGGAGCGTCCGCGAGGGGAGACCTG 634
 QY 181 GluGlnIleLeuThrGlnTyrThrThrPheVal 191
 DB 635 GAGCAGATTCTGACGAGTACACACCTTCGTG 667

RESULT 7

AAH75355 ID AAH75355 standard; cDNA; 900 BP.

AC AAH75355;

DT 02-OCT-2001 (first entry)

XX Human uridine kinase encoding cDNA.

XX Human; uridine kinase; UK; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

FT CDS 8..838

FT /product= "uridine kinase"

FT /note= "Claimed in claim 1"

XX CN1287172-A.

XX 14-MAR-2001.

XX PF 07-SEP-1999; 99CN-0118818.

XX PR 07-SEP-1999; 99CN-0118818.

XX PA (UYFU-) UNIV FUDAN.

XX PI Yu L, Zhao Y, Zhang H;

XX DR WPI; 2001-409529/44.

XX DR P-PSDB; AAG64506.

XX PT Human uridine kinase and its coding sequence, preparation and

XX application -

XX PS Claim 1; Page 15(Disclosure); 20pp; Chinese.

XX CC The invention relates to human uridine kinase (UK).

XX SQ Sequence 900 BP; 201 A; 237 C; 292 G; 170 T; 0 other;

Alignment Scores:

Pred. No.: 1e-180 Length: 900
 Score: 189.00 Matches: 258
 Percent Similarity: 98.47% Conservative: 0
 Best Local Similarity: 98.47% Mismatches: 2
 Query Match: 68.23% Indels: 4
 DB: 22 Gaps: 0

US-09-896-522-2 (1-277) x AAH75355 (1-900)

QY 18 ArgProHisGlnIleArgProPheLeuIleGlyValSerGlyGlyThrAlaSerGlyLysSer 37
 DB 56 CGTCCGACCAAGGTCCTTCCTGATAGGGGTAGGGCGGCACTGCCACGGGAAAGTCG 115
 QY 38 ThrValCysGlnIleMetGluLeuGluGlyGlnAsnGluValGluGlnArgGlnArg 57
 DB 116 ACCGTGTGAGAAAGATCATGAGATTGCTGGGACAGAACAGAGGTGAAACCGGCA-CGG 174
 QY 58 -LysValValIleLeuSerGlnAspArgPheTyrLysValLeuThrAlaGluGlnLysAl 77
 DB 175 TAAAGTGTATCTGATCGACAGGAGTTCATCAAGTCTCTGACGCGACAGAGAGAGGC 234
 QY 77 alyValAlaLeuLysGlyGlnTyrAsnPheAspHisProAspAlaPheAspAsnAspLeuMe 97
 DB 235 CAAGGCCCTTGAAAGGACAGTAAATTGTCATTCAGAGTCCCTTGATATATATTGAT 294
 QY 97 tHisArgThrLeuValAsnIleValGluGlyLysThrValGluValProThrTyrAspPh 117
 DB 295 GCAAGGACTCTTAAGAACATCGTGGAGGCAAAACGGTGGAGGCGACCTATGATTT 354
 QY 117 eValThrHisSerArgLeuProGluThrThrValValTyrProAlaAspValValLeuPh 137
 DB 355 TGTGACACACTCAAGGTTACAGAGACCAAGGTCGTACCTCGGAGCGTGTCTGTT 414
 QY 137 eGluGlyIleLeuValPheTyrSerGluGluIleArgAspMetPheHisLeuArgLeuPh 157
 DB 415 TGAAGGATCTTGTGTCTACAGCCAGAGAGATCCGGGACATGTTCCACTGCGCTCTT 474
 QY 157 eValAspThrAspSerAspValArgLeuSerArgValLeuArgAspValArgArgGly 177
 DB 475 CGTGGACACCGACTCCGAGCGTCAGGCTGTCTCAAAAGTTCCTCGGAGCGTGGCCGG 534
 QY 177 YArgAspLeuGluGlnIleLeuThrGlnTyrThrThrPheValLysProAlaPheGluGly 197
 DB 535 GAGGAGACTGGACAGATTCAGCGACAGTACACACCTTCGTGAGCGGCTTCGAGGA 594
 QY 197 uPheCysLeuProThrLysLysTyrAlaAspValIleIleProArgGlyValAspAsnMe 217
 DB 595 GTTCTGCTCCGCCACAAAGAGTATCCGATGATCATCCACGAGGTGACAAATAT 654
 QY 217 tValAlaIleAsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysLys 237
 DB 655 GGTTCGATCACTGATCTGCGACGACATCCAGACATTCGAAAGTGACATTCGCAA 714
 QY 237 sTrpHisArgGlyGlySerAsnGlyArgSer-TyrLysArgThrPheSerGluProGlyA 257
 DB 715 ATGGACCGAGGAGGATCCATATGGCGGAC-GTACAAAGCGGACCTTTCTGAGCCAGGGG 773
 QY 257 sPheIleProGlyMetLeuThrSerGlyLysArgSerHisLeuGluSerSerArgProH 277
 DB 774 ACCACCTGGAGATGATCTGCGCAAAACGGTCACATTTGGAGTCCAGACAGACACCCC 833
 QY 277 is 277

DB 834 AC 835

RESULT 8

ABL89762 ID ABL89762 standard; cDNA; 1396 BP.

AC ABL89762;

DT 24-MAY-2002 (first entry)

XX Human polynucleotide SEQ ID NO 324.

DE Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
 XX antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor;
 XX valnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
 KM cardiatic; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KM neurological disease; infection; human; secreted protein; gene; ss.

XX OS Homo sapiens.
 XX PN WO200190304-A2.
 XX PD 29-NOV-2001.
 XX PF 18-MAY-2001; 2001WO-US16450.
 XX PR 19-MAY-2000; 2000US-205515P.
 XX PA (HUMA-) HUMAN GENOME SCI INC.
 XX PI Birse CE, Rosen CA;
 XX DR WPI; 2002-122018/16.
 XX DR P-PSDB; ABB89353.
 XX PT Novel 1405 isolated polypeptides, useful for diagnosis, treatment and
 PT prevention of neural, immune system, muscular, reproductive,
 PT gastrointestinal, pulmonary, cardiovascular, renal and proliferative
 PT disorders -
 XX
 XX PS Claim 4; SEQ ID NO 324; 2081bp + Sequence Listing; English.
 XX
 XX CC The invention relates to novel genes (ABU89449-ABU90853) and proteins
 CC (ABB890444) useful for preventing, treating or ameliorating
 CC medical conditions e.g. by protein or gene therapy. The genes are
 CC isolated from a range of human tissues disclosed in the specification.
 CC The nucleic acids, proteins, antibodies and (ant)agonists are useful
 CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
 CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
 CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
 CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
 CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
 CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
 CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;
 CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
 CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
 CC and parasitic infections.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 XX
 XX SQ Sequence 1396 BP; 317 A; 358 C; 425 G; 287 T; 9 other;
 Alignment Scores:
 Pred. No.: 3,84e-127 Length: 1396
 Score: 136.00 Matches: 244
 Percent Similarity: 97.99% Conservative: 0
 Best Local Similarity: 97.99% Mismatches: 3
 Query Match: 49.10% Indels: 5
 DB: 24 Gaps: 0
 US-09-896-522-2 (1-277) x ABI89762 (1-1396)
 QY 1 MetAlaSerAlaGlyGlyLysAspCysGlySerProAlaProGluAlaAspArgProHis 20
 Db 41 ATGGCTTGGCGGCGAGCGAGACTGCGAGAGCCCCCGCGAGCGCCAGCTCCGAC 100
 QY 21 GlnArgProPheLeuIleGlyValSerGlyGlyThraIa-SerGlyysSerThraIaCy 40
 Db 101 CAGCGGCGCCCTTCCTGATAGGGGTGAGCGCGGCACTG-CAGCGGGAAGTGCAGCGGTG 159
 QY 40 sGluLysIleMetGluLeuLeuGlyGlnAsnGluValGlnGlnArgGlnArgGlyValVa 60
 Db 160 TGAAGAAGTCAATGAGTGTGCGGACAGAAAGAGGTGMAACGCGCGCAGGAAGTGTGT 219
 QY 111LeuSerGlnAspArgPheTyrLysValIleuThraIaGlnGlnLysAlaIysAlaIe 80
 Db 220 CATCTGAGCGAGAGAGGTTCTTACAGGCTCTTACGCGCAGAGCAAGAGGCTTT 279
 QY 80 uLysGlyGlnTyrAsnPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgTh 100

Db 280 GAAAGACAGTACAAATTTTATGATCCAGATGCTTTGATTAATGATTGACAGAGAC 339
 QY 100 rLeuLysAsnIleValGlnGlySerThraValGluValProThrTyrAspPheValThrH 120
 Db 340 TCTGAAGAACATCGTGAAGGGCAAAACGGTGAAGGTGCCAGCTTAATGATTTTGTACACA 399
 QY 120 sSerArgLeuProGluThrThraValTyrProAlaAspValValLeuPheGluGlyI 140
 Db 400 CTCAAGGTTACAGAGACCCAGGTGTCTACCTTGGCGAGCTGTCTGTTAAGGGCAT 459
 QY 140 eLeuValPheTyrSerGlnGlnIleArgAspMetPheHisLeuArgLeuPheValAspTh 160
 Db 460 CTGTGTCTTACAGCAGAGATCCGGGACATGTTCCACTCGCCTCTTGTGACAC 519
 QY 160 rAspSerAspValArgLeuSerArgArgVal-LeuArgAspValArgArgGlyArgAspL 180
 Db 520 CGACTCCGACGTCAGGTGTCTGAAGAAKT-TCTCCGGACGTGCGCGGAGGAGGAC 578
 QY 180 euGlnGlnIleLeuThrGlnTyrThrThraPheValLysProAlaPheGluGluPheCysL 200
 Db 579 TGAAGCAGATTCTGACCCA-TACACCACTTCTGTAAGCCGCTTCGAGGATTTCTGCC 637
 QY 200 euProThrLysLysTyrAlaAspValIleIleProArgGlyValAspAsnMetValAlaI 220
 Db 638 TGCAGCAAGAGATGCGGATGTGATCATCCACGAGAGAGTGACAAATATGTTGCCA 697
 QY 220 LeuAsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysLysTyrHisA 240
 Db 698 TCAACCGATCGTGCGACGACATCCAGACATTCTGAATGGTGACATCTGAATGGCACCC 757
 QY 240 rGlyGlySerAsnGlyArgSer 247
 Db 758 GAGGAGGTCCTCAATGGCGGAGT 760
 RESULT 9
 ABSS1801
 ID ABSS1801 standard; cDNA; 734 BP.
 XX
 AC ABSS1801;
 XX
 DT 21-OCT-2002 (first entry)
 XX
 DE Human mddt cDNA Incyte ID No: LI:235557.12:2001JAN12.
 XX
 KW Human; molecule for disease detection and treatment; MDDT; cancer;
 KW cell proliferative disorder; arteriosclerosis; cirrhosis; asthma;
 KW autoimmune disorder; inflammatory disorder; Crohn's disease;
 KW multiple sclerosis; cytostatic; antiarteriosclerotic; antiinflammatory;
 KW hepatotropic; immunosuppressive; antiasthmatic; gene therapy; gene; ss.
 OS Homo sapiens.
 XX
 XX PN WO200255738-A2.
 XX PD 18-JUL-2002.
 XX PF 09-JAN-2002; 2002WO-US01008.
 XX
 XX 12-JAN-2001; 2001US-261622P.
 XX 16-JAN-2001; 2001US-261865P.
 XX 17-JAN-2001; 2001US-262208P.
 XX 17-JAN-2001; 2001US-262209P.
 XX 17-JAN-2001; 2001US-262326P.
 XX 17-JAN-2001; 2001US-262326P.
 XX 19-JAN-2001; 2001US-263063P.
 XX 19-JAN-2001; 2001US-263065P.
 XX 19-JAN-2001; 2001US-263329P.
 XX
 XX PA (INCYTE GENOMICS INC.
 XX Panzer SR, Lincoln SE, Altus CM, Dufour GE, Hillman JL, Jones AL,
 XX Dem TC, Liu TF, Harris B, Flores V, Dafo A, Marwaha R, Chen AD;
 XX Chang SC, Gerstein EH, Peralta CH, David MH, Lewis SA;

XX WPI: 2002-590679/63.
DR P-SDB; ABG70328.
XX
PT New disease detection and treatment molecule (MDT) polynucleotides and
PT polypeptides, useful in diagnosing, studying, preventing or treating
PT diseases associated with MDT expression, e.g. autoimmune or
PT inflammatory disorders
XX
PS Claim 1; Page 98; 129pp; English.
XX
CC The present invention relates to the isolation of novel human
CC molecules for disease detection and treatment (MDT), and the
CC polynucleotide sequences (mdt) encoding them. The MDT polypeptides
CC may be used to screen for molecules that bind to, or are bound by the
CC encoded polypeptides, and to develop a transcript image of a tissue or
CC cell type. Probes comprising at least 20 nucleotides of the mdt
CC polynucleotide may be used to assess the toxicity of a test compound.
CC The MDT polypeptides and mdt polynucleotides are useful in the
CC diagnosis, study, prevention and treatment of diseases associated with
CC the expression of molecules for disease detection and treatment. Such
CC disorders include cell proliferative disorders (e.g. arteriosclerosis,
CC cirrhosis, or cancers), and autoimmune/inflammatory disorders
CC (e.g. asthma, Crohn's disease, or multiple sclerosis). The mdt
CC polynucleotides may also be used as molecule markers, in microarrays,
CC and in somatic or germline gene therapy. ABS51779-ABS51814 encode
CC the MDT proteins of the invention.
XX
SQ Sequence 734 BP; 166 A; 181 C; 226 G; 154 T; 7 other;
Alignment Scores:
Pred. No.: 3.21e-103 Length: 734
Score: 112.00 Matches: 112
Percent Similarity: 100.00% Conserves: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 40.43% Indels: 0
Gaps: 0
US-09-896-522-2 (1-277) x ABS51801 (1-734)
QY 90 AspAlaPheAspAspLeuMetHisArgThrLeuLysAsnIleValIleGlyLysThr 109
Db 14 GATGCTTTGATATGATTGATGACACAGACTCTGAAGAATCGTGAAGGCAAAACG 73
QY 110 ValIleValProThrTyrAspPheValThiHisSerArgLeuProGluThrValVal 129
Db 74 GTGAGGAGTCCGACCTATGATTGTGACACACCAAGTATACAGAGACCAACGTCGTC 133
QY 130 TyrProAlaAspValValIleuPheGluGlyIleLeuValPheTyrSerGlnIleArg 149
Db 134 TACCTCGGACGCGGTTCCTTTGAGGGCATCTTGTGTTCTTACACCGAGATCCGG 193
QY 150 AspMetPheHisLeuArgLeuPheValAspThrAspSerAspValArgLeuSerArgArg 169
Db 194 GACATGTTCCACGCGCTCTTCTGTGACACGACCTCCGACGTCGAGTCTCGAAGA 253
QY 170 ValLeuArgAspValArgArgGlyArgAspLeuGluGlnIleLeuThrGlnTyrThrThr 189
Db 254 GTTCTCGGAGCGCGCCGACGAGGAGGACCTGAGGACGATTTCAGACGATACACACC 313
QY 190 PheValIysProAlaPheGluGluPheCysLeuPro 201
Db 314 TTCGTGAGCCGCGCTTCGAGAGATTCTGCTGCGG 349
RESULT 10
AAL36439
ID AAL36439 standard; DNA; 9732 BP.
XX
AC AAL36439;
XX
DT 08-JAN-2002 (first entry)
XX
DE Human musculoskeletal system related polynucleotide SEQ ID NO 2804.

XX
KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; anticancer;
KW vlnenarary; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW neurological disease; infection; human; secreted protein;
KW musculoskeletal system; ds.
XX
OS Homo sapiens.
XX
PN WO200155367-A1.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01338.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220863.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226688.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.

```
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234222.
PR 21-SEP-2000; 2000US-0234276.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235835.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-024617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.

PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-451937/48.
XX
XX Isolated polypeptide for treating, preventing and/ or prognosing
XX disorders related to the musculoskeletal system including
XX musculoskeletal cancers and also for testing and detection e.g.
XX diagnosis -
XX
XX Example 2; SEQ ID NO 2804; 781pp + Sequence Listing; English.
XX
XX The invention relates to novel genes (AAL34669-AAL37666) and proteins
XX (ABB01087-ABB04109) associated with the musculoskeletal system useful
XX for preventing, treating or ameliorating medical conditions e.g. by
XX protein or gene therapy. The genes are isolated from a range of human
XX tissues disclosed in the specification. The nucleic acids, proteins,
XX antibodies and (ant)agonists are useful in the diagnosis, treatment
XX and prevention of: (a) cancer, e.g. breast and ovarian cancer and
XX other cancers of the adrenal gland, bone, bone marrow, breast,
XX gastrointestinal tract, liver, lung, or urogenital; (b) immune
XX disorders e.g. Addison's disease, allergies, autoimmune hemolytic
XX anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
XX multiple sclerosis, rheumatoid arthritis and ulcerative colitis;
XX (c) cardiovascular disorders such as myocardial ischaemia; (d) wound
XX healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy;
XX and (f) infectious diseases such as viral, bacterial, fungal and
XX parasitic infections.
XX
XX Note: The sequence data for this patent did not form part of the
XX printed specification, but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 9732 BP; 2063 A; 2805 C; 2561 G; 2303 T; 0 other;
XX
XX
XX Alignment Scores:
XX Pred. No.: 8.93e-50 Length: 9732
XX Score: 60.00 Matches: 60
XX Percent Similarity: 100.00% Conservative: 0
XX Best Local Similarity: 100.00% Mismatches: 0
XX Query Match: 21.66% Indels: 0
XX DB: 22 Gaps: 0
XX
XX US-09-896-522-2 (1-277) x AAL36439 (1-9732)
XX
XX QY 218 ValAlaIleAsnLeuIleValGlnHisIleGlaAspIleLeuAsnGlyAspIleCysIle 237
XX Db 3901 GTTGCATCAACCTGATCGGACGACATCCAGACATTTCTGATGTCATCTCCAA 3960
XX QY 238 TrpHisArgGlyGlySerAsnGlyArgSerTyrIleArgIlePheSerGluProGlyAsp 257
XX Db 3961 TGGACCGAGAGAGGTCACATGGGCGAGCTTACAGCGGACCTTTCTGAGCCAGGAGAC 4020
XX QY 258 HisProGlyMetLeuThrSerGlyIleArgSerHisLeuGluSerSerArgProHis 277
XX Db 4021 CACCTGGGATGCTGACCTTGGCAACGCTGACATTTGAGATCCAGACAGCCACC 4080
XX
XX RESULT 11
XX ABX59427
XX ID ABX59427 standard; cDNA; 9732 BP.
XX
XX AC ABX59427;
XX
XX DT 26-FEB-2003 (first entry)
XX
XX DE cDNA encoding novel human musculoskeletal system antigen #1771.
```

XX Gene; ss; musculoskeletal system antigen; cancer; metastasis;
KM re-vascularisation; thrombosis; arteriosclerosis; mineral content;
KM cardiovascular condition; wound; injury; burn; angiogenesis; ulcer;
KM post-operative tissue repair; limb regeneration; neuronal growth;
KM neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
KM AIDS-related complex; chondrocyte growth; bone regeneration;
KM periodontal regeneration; tissue transport; bone graft; skin aging;
KM keratinocyte growth; hair loss; melanocyte growth; cell proliferation;
KM cell growth; organ transplant; cell differentiation; body height;
KM weight; hair colour; eye colour; skin; percentage of adipose tissue;
KM pigmentation; cosmetic surgery; metabolism; biorhythm; cardiac rhythm;
KM depression; tendency for violence; pain; reproductive capability;
KM hormone level; endocrine level; appetite; libido; memory; stress;
KM storage capability; fat content; lipid content; protein content;
KM carbonate content; vitamin content; cofactor content;
XX nutritional component.
XX Homo sapiens.
XX OS
XX US2002147140-A1.
XX PN
XX 10-OCT-2002.
XX PD
XX 17-JAN-2001; 2001US-0764877.
XX PF
XX 31-JAN-2000; 2000US-179065P.
PR 04-FEB-2000; 2000US-180628P.
PR 28-JUN-2000; 2000US-214886P.
PR 07-JUL-2000; 2000US-216647P.
PR 07-JUL-2000; 2000US-216880P.
PR 11-JUL-2000; 2000US-217487P.
PR 11-JUL-2000; 2000US-217496P.
PR 14-JUL-2000; 2000US-218290P.
PR 26-JUL-2000; 2000US-220963P.
PR 26-JUL-2000; 2000US-220964P.
PR 14-AUG-2000; 2000US-224518P.
PR 14-AUG-2000; 2000US-224519P.
PR 14-AUG-2000; 2000US-225267P.
PR 14-AUG-2000; 2000US-225268P.
PR 14-AUG-2000; 2000US-225270P.
PR 14-AUG-2000; 2000US-225447P.
PR 14-AUG-2000; 2000US-225757P.
PR 14-AUG-2000; 2000US-225758P.
PR 22-AUG-2000; 2000US-226868P.
PR 30-AUG-2000; 2000US-228924P.
PR 01-SEP-2000; 2000US-229287P.
PR 01-SEP-2000; 2000US-229343P.
PR 01-SEP-2000; 2000US-229344P.
PR 01-SEP-2000; 2000US-229345P.
PR 05-SEP-2000; 2000US-229509P.
PR 05-SEP-2000; 2000US-229513P.
PR 08-SEP-2000; 2000US-231413P.
PR 21-SEP-2000; 2000US-234223P.
PR 21-SEP-2000; 2000US-234274P.
PR 25-SEP-2000; 2000US-234977P.
PR 27-SEP-2000; 2000US-235834P.
PR 29-SEP-2000; 2000US-236327P.
PR 29-SEP-2000; 2000US-236367P.
PR 29-SEP-2000; 2000US-236368P.
PR 29-SEP-2000; 2000US-236369P.
PR 29-SEP-2000; 2000US-236370P.
PR 02-OCT-2000; 2000US-236802P.
PR 02-OCT-2000; 2000US-237037P.
PR 02-OCT-2000; 2000US-237038P.
PR 02-OCT-2000; 2000US-237039P.
PR 02-OCT-2000; 2000US-237040P.
PR 13-OCT-2000; 2000US-239335P.
PR 20-OCT-2000; 2000US-240960P.
PR 20-OCT-2000; 2000US-241785P.
PR 20-OCT-2000; 2000US-241809P.
PR 01-NOV-2000; 2000US-244617P.
PR 17-NOV-2000; 2000US-249299P.

PR 08-DEC-2000; 2000US-251856P.
PR 08-DEC-2000; 2000US-251868P.
PR 08-DEC-2000; 2000US-251869P.
XX
PA (ROSE/) ROSEN C A.
PA (RUBE/) RUBEN S M.
PA (BARA/) BARASH S C.
XX
PI Rosen CA, Ruben SM, Barash SC,
XX WPI; 2003-128199/12.
XX
XX
XX Isolated nucleic acid molecules encoding musculoskeletal system
PT associated polypeptides, useful for detecting disorders, e.g. cancer -
PS disclosure, SEQ ID NO 2804; 321pp; English.
XX
XX The invention describes an isolated nucleic acid molecule comprising a
CC sequence encoding musculoskeletal system associated polypeptides useful
CC for detecting disorders, e.g., cancer or cancer metastases, in animals
CC or humans. The nucleic acid stimulates re-vascularisation of ischaemic
CC tissues associated with conditions such as thrombosis, arteriosclerosis,
CC and other cardiovascular conditions; treats wounds due to injuries,
CC burns, post-operative tissue repair, and ulcers; stimulates angiogenesis
CC and limb regeneration; stimulates neuronal growth; can treat and prevent
CC neuronal damage occurring in certain disorders or neurodegenerative
CC conditions, such as, Alzheimer's disease, Parkinson's disease, and
CC AIDS-related complex; stimulates chondrocyte growth, thus they can be
CC used to enhance bone and periodontal regeneration and aid in tissue
CC transports or bone grafts; prevents skin aging due to sunburn by
CC stimulating keratinocyte growth; prevents hair loss, since FGF family
CC members activate hair-forming cells and promotes melanocyte growth;
CC stimulates growth and differentiation of hematopoietic cells and bone
CC marrow cells when used in combination with other cytokines; maintains
CC organs before transplantation or for supporting cell culture of primary
CC tissues; induces tissue of mesodermal origin to differentiate in early
CC embryos; increases or decreases the differentiation or proliferation of
CC embryonic stem cells, besides, haematopoietic lineage; modulates
CC mammalian characteristics, such as, body height, weight, hair colour, eye
CC colour, skin, percentage of adipose tissue, pigmentation, size, and shape
CC (e.g., cosmetic surgery); modulates mammalian metabolism; changes
CC mammal's metabolic state or physical state by influencing biorhythms,
CC cardiac rhythms, depression, tendency for violence, tolerance for pain,
CC reproductive capabilities, hormonal or endocrine levels, appetite,
CC libido, memory, or stress; increases or decreases storage capabilities,
CC fat content, lipid, protein, carbohydrate, vitamins, minerals, cofactors
CC or other nutritional components. This sequence encodes a novel human
CC musculoskeletal system antigen.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from the US patent office at
CC ftp.segdata.uspto.gov/sequence.html?DocID=20020147140.
XX
XX Sequence 9732 BP; 2063 A; 2805 C; 2561 G; 2303 T; 0 other;
SQ
XX
XX Alignment Scores:
Pred. No.: 8,93e-50 Length: 9732
Score: 60.00 Matches: 60
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 21.66% Indels: 0
DB: 25 Gaps: 0
US-09-896-522-2 (1-277) x ABX59427 (1-9732)
QY 218 ValAlaIleAsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysLys 237
Db 3901 GTTGCATCACTAGTCTGTCGACGACATTCGAGCATTCGATCTGCAAA 3960
QY 238 TTPHISARGGlyGlySerAsnGlyArgSerTyrLysArgThrPheSerGluProGlyAsp 257
Db 3961 TGGCACCGAGGAGGCTCAATGGGCGGAGCTCAAGACGCGACTTTTCTGACCGAGGAGAC 4020

Oy 258 HisProGlyMetLeuThrSerGlyLysArgSerHisLeuGluSerSerArgProHis 277
Db |||||
4021 CACCTGGAGTCTGACTCTGGCAACGTCACATTGGAGTCCAGCAGACCCAC 4080
RESULT 12
AAL36440
ID AAL36440 standard; DNA; 19125 BP.
XX
AC AAL36440;
XX
DT 08-JAN-2002 (first entry)
XX
DE Human musculoskeletal system related polynucleotide SEQ ID NO 2805.
XX
XX Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;
KW vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW neurological disease; infection; human; secreted protein;
musculoskeletal system; ds.
XX
XX Homo sapiens.
XX
XX MO20015367-A1.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01338.
XX
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.

PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234597.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.

	PR	17-NOV-2000;	2000US-0249245.
	PR	17-NOV-2000;	2000US-0249264.
	PR	17-NOV-2000;	2000US-0249265.
	PR	17-NOV-2000;	2000US-0249297.
	PR	17-NOV-2000;	2000US-0249299.
	PR	17-NOV-2000;	2000US-0249300.
	PR	01-DEC-2000;	2000US-0250160.
	PR	01-DEC-2000;	2000US-0250391.
	PR	05-DEC-2000;	2000US-0251030.
	PR	05-DEC-2000;	2000US-0251988.
	PR	05-DEC-2000;	2000US-0256719.
	PR	06-DEC-2000;	2000US-0251479.
	PR	08-DEC-2000;	2000US-0251856.
	PR	08-DEC-2000;	2000US-0251868.
	PR	08-DEC-2000;	2000US-0251869.
	PR	08-DEC-2000;	2000US-0251989.
	PR	11-DEC-2000;	2000US-0254097.
	PR	05-JAN-2001;	2001US-0259678.
	XX	(HUMA-) HUMAN GENOME SCI INC.	
	PA	Rosen CA, Barash SC, Ruben SM;	
	P1	WPI; 2001-451937/48.	
	XX		
	DR		
	PT	Isolated polypeptide for treating, preventing and/or prognosing	
	PT	disorders related to the musculoskeletal system including	
	PT	musculoskeletal cancers and also for testing and detection e.g.	
	PT	diagnosis -	
	PS	Example 2; SEQ ID NO 2805; 781pp + Sequence Listing; English.	
	XX		
	CC	The invention relates to novel genes (AAI34669-AAJ37666) and proteins	
	CC	(AAB03087-ABB04109) associated with the musculoskeletal system useful	
	CC	for preventing, treating or ameliorating medical conditions e.g. by	
	CC	protein or gene therapy. The genes are isolated from a range of human	
	CC	tissues disclosed in the specification. The nucleic acids, proteins,	
	CC	antibodies and (anti)agonists are useful in the diagnosis, treatment	
	CC	and prevention of: (a) cancer, e.g. breast and ovarian cancer and	
	CC	other cancers of the adrenal gland, bone, bone marrow, breast,	
	CC	gastrointestinal tract, liver, lung, or urogenital; (b) immune	
	CC	disorders e.g. Addison's disease, allergies, autoimmune haemolytic	
	CC	anemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,	
	CC	multiple sclerosis, rheumatoid arthritis and ulcerative colitis;	
	CC	(c) cardiovascular disorders such as myocardial ischaemias; (d) wound	
	CC	healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy;	
	CC	and (f) infectious diseases such as viral, bacterial, fungal and	
	CC	parasitic infections.	
	CC	Note: The sequence data for this patent did not form part of the	
	CC	printed specification, but was obtained in electronic format directly	
	CC	from WIPO at ftp.wipo.int/pub/published_pct_sequences.	
	XX		
	SQ	Sequence 19125 BP; 4263 A; 5265 C; 5444 G; 4153 T; 0 other;	
	Alignment Scores:		
	Pred. No.:	1.68e-49	Length: 19125
	Scores:	60.00	Matches: 60
	Percent Similarity:	100.00%	Conservative: 0
	Best Local Similarity:	100.00%	Mismatches: 0
	Query Match:	21.66%	Indels: 0
	DB:	22	Gaps: 0
	US-09-896-522-2 (1-277) x AAI36440 (1-19125)		
OY	218	VAlAAIAleAsnLeuileValGlnHisillegInApIleleuASnglyASpIleCylys	237
Db	5997	GTTGCACCAACCTGATCGTCAGACATCCAGGACATTGTGAATGGTGAATCTGCCAA	6056
OY	238	TThpIsArGglYglYlSeASnglyARgeerYrlylaAgTrRpHeSeGlUpProGlyASP	257
Db	6057	TGGACACCAAGAGGGTCCAAATGGGGGAGCTPACAAGCGAGACTTTTGAGCCAGGGGAC	6116

Db	Accession	Gene	Protein	Length	Start	End	Strand	Score	E-value	Annotations
Oy	258	HistH3.1	HistH3.1	277	1	277	+	100.0	0.0	Gene, histone H3.1
Db	6117	CACCTGGATCTGACCTCTGGGAAAGCGTCACATTGGAGTCCAGCAGACCCGAC		6178	1	6178	+	100.0	0.0	cdna encoding novel human musculoskeletal system antigen #1772.
RESULT 13										
ID	ABX59428	ABX59428 standard; cdna; 19125 BP.								
XX	AC	ABX59428;								
XX	DT	26-FEB-2003 (first entry)								
XX	DE	cdna encoding novel human musculoskeletal system antigen #1772.								
KM	KM	Gene, ss; musculoskeletal system antigen; cancer; metastasis;								
KM	KM	re-vascularisation; thrombosis; arteriosclerosis; mineral content;								
KM	KM	cardiovascular condition; wound; injury; burn; angiogenesis; ulcer;								
KM	KM	post-operative tissue repair; limb regeneration; neuronal growth;								
KM	KM	neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;								
KM	KM	AIDS-related complex; chondrocyte growth; bone regeneration;								
KM	KM	periodontal regeneration; tissue transport; bone graft; skin aging;								
KM	KM	keratinocyte growth; hair loss; melanocyte growth; cell proliferation;								
KM	KM	cell growth; organ transplant; cell differentiation; body height;								
KM	KM	weight; hair colour; eye colour; skin; percentage of adipose tissue;								
KM	KM	pigmentation; cosmetic surgery; metabolism; biorhythm; cardiac rhythm;								
KM	KM	depression; tendency for violence; pain; reproductive capability;								
KM	KM	hormone level; endocrine level; appetite; libido; memory; stress;								
KM	KM	storage capability; fat content; lipid content; protein content;								
KM	KM	carbohydrate content; vitamin content; cofactor content;								
XX	XX	nutritional component.								
XX	OS	Homo sapiens.								
XX	PN	US2002147140-A1.								
XX	PD	10-OCT-2002.								
XX	PF	17-JAN-2001; 2001US-0764877.								
XX	PR	31-JAN-2000; 2000US-179065P.								
XX	PR	04-FEB-2000; 2000US-180628P.								
XX	PR	28-JUN-2000; 2000US-214886P.								
XX	PR	07-JUL-2000; 2000US-216647P.								
XX	PR	07-JUL-2000; 2000US-216880P.								
XX	PR	11-JUL-2000; 2000US-217487P.								
XX	PR	11-JUL-2000; 2000US-217496P.								
XX	PR	14-JUL-2000; 2000US-218290P.								
XX	PR	26-JUL-2000; 2000US-220963P.								
XX	PR	26-JUL-2000; 2000US-220964P.								
XX	PR	14-AUG-2000; 2000US-224518P.								
XX	PR	14-AUG-2000; 2000US-224519P.								
XX	PR	14-AUG-2000; 2000US-225267P.								
XX	PR	14-AUG-2000; 2000US-225268P.								
XX	PR	14-AUG-2000; 2000US-225270P.								
XX	PR	14-AUG-2000; 2000US-225477P.								
XX	PR	14-AUG-2000; 2000US-225757P.								
XX	PR	14-AUG-2000; 2000US-225758P.								
XX	PR	22-AUG-2000; 2000US-226868P.								
XX	PR	30-AUG-2000; 2000US-228924P.								
XX	PR	01-SEP-2000; 2000US-228924P.								
XX	PR	01-SEP-2000; 2000US-229343P.								
XX	PR	01-SEP-2000; 2000US-229344P.								
XX	PR	01-SEP-2000; 2000US-229345P.								
XX	PR	05-SEP-2000; 2000US-229509P.								
XX	PR	05-SEP-2000; 2000US-229513P.								
XX	PR	08-SEP-2000; 2000US-231413P.								

PR 29-SEP-2000; 2000US-236369P.
 PR 29-SEP-2000; 2000US-236370P.
 PR 02-OCT-2000; 2000US-236802P.
 PR 02-OCT-2000; 2000US-237037P.
 PR 02-OCT-2000; 2000US-237038P.
 PR 02-OCT-2000; 2000US-237039P.
 PR 13-OCT-2000; 2000US-237040P.
 PR 20-OCT-2000; 2000US-240960P.
 PR 20-OCT-2000; 2000US-241785P.
 PR 20-OCT-2000; 2000US-241809P.
 PR 01-NOV-2000; 2000US-244617P.
 PR 17-NOV-2000; 2000US-249299P.
 PR 08-DEC-2000; 2000US-251856P.
 PR 08-DEC-2000; 2000US-251868P.
 PR 08-DEC-2000; 2000US-251869P.
 PR 08-DEC-2000; 2000US-251869P.
 PA (ROSE/) ROSEN C A.
 PA (RUBE/) RUBEN S M.
 PA (BARA/) BARASH S C.
 XX
 PI Rosen CA, Ruben SM, Barash SC;
 XX
 DR WPI: 2003-128199/12.
 XX
 PT Isolated nucleic acid molecules encoding musculoskeletal system
 XX associated polypeptides, useful for detecting disorders, e.g. cancer -
 XX
 PS Disclosure; SEQ ID NO 2805; 31pp; English.
 XX
 CC The invention describes an isolated nucleic acid molecule comprising a
 CC sequence encoding musculoskeletal system associated polypeptides useful
 CC for detecting disorders, e.g., cancer or cancer metastases, in animals
 CC or humans. The nucleic acid: stimulates re-vascularization of ischemic
 CC tissues associated with conditions such as thrombosis, arteriosclerosis,
 CC burns, post-operative tissue repair, and ulcers; stimulates angiogenesis
 CC and limb regeneration; stimulates neuronal growth; can treat and prevent
 CC neuronal damage occurring in certain disorders or neurodegenerative
 CC conditions, such as, Alzheimer's disease, Parkinson's disease, and
 CC AIDS-related complex; stimulates chondrocyte growth, thus they can be
 CC used to enhance bone and periodontal regeneration and aid in tissue
 CC transports or bone grafts; prevents skin aging due to sunburn by
 CC stimulating keratinocyte growth; prevents hair loss, since FGF family
 CC members activate hair-forming cells and promotes melanocyte growth;
 CC stimulates growth and differentiation of hematopoietic cells and bone
 CC marrow cells when used in combination with other cytokines; maintains
 CC organs before transplantation or for supporting cell culture of primary
 CC tissues; induces tissue of mesodermal origin to differentiate in early
 CC embryos; increases or decreases the differentiation or proliferation of
 CC embryonic stem cells, besides, hematopoietic lineage; modulates
 CC mammalian characteristics, such as, body height, weight, hair colour, eye
 CC colour, skin, percentage of adipose tissue, pigmentation, size, and shape
 CC (e.g., cosmetic surgery); modulates mammalian metabolism, changes
 CC mammal's metal state or physical state by influencing biorythms,
 CC cardiac rhythms, depression, tendency for violence, tolerance for pain,
 CC reproductive capabilities, hormonal or endocrine levels, appetite,
 CC libido, memory, or stress; increases or decreases storage capabilities,
 CC fat content, lipid, protein, carbohydrate, vitamins, minerals, cofactors
 CC or other nutritional components. This sequence encodes a novel human
 CC musculoskeletal system antigen.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from the US patent office at
 CC ftp.seqdata.uspto.gov/sequence.html?DocID=20020147140.
 XX
 SQ Sequence 19125 BP; 4263 A; 5265 C; 5444 G; 4153 T; 0 other;
 Alignment Scores:
 Pred. No.: 1,68e-49 Length: 19125
 Score: 60.00 Matches: 60
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0

Query Match: 21.66% Indels: 0
 DB: 25 Gaps: 0
 US-09-896-522-2 (1-277) x ABX59428 (1-19125)
 QY 218 ValAlaIleAsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysLys 237
 Db 5997 GTTGCCATCAACCTGATCGTGCAGCAGCATCCAGACACTTCTGAATGGTGAACCTGCAG 6056
 QY 238 TrpHisArgGlyGlySerAsnGlyArgSerTyrLysArgThrPheSerGluPProGlyAsp 257
 Db 6057 TGGCACCAGAGAGGCTCAATGGGCGAGCTAACAGCGACCTTTTCTGAGCCAGGGGAC 6116
 QY 258 HisProGlyMetLeuThrSerGlyLysArgSerHisLeuGluIleSerSerArgProHis 277
 Db 6117 CACCTGGAGTGGTGAACCTTGGCAACGGATCATTGGAGTTCAGACAGACAGCCCCAC 6176
 RESULT 14
 AA215454
 ID AA215454 standard; cDNA; 715 BP.
 XX
 AC AA215454;
 XX
 DT 12-OCT-1999 (first entry)
 XX
 DE Human gene expression product cDNA sequence SEQ ID NO:2923.
 XX
 KW Human; gene; gene expression product; diagnosis; therapy; probe;
 KW detection; mapping; tissue typing; profiling; forensic; cancer;
 KW genetic analysis; colorectal cancer; breast cancer; lung cancer; ss.
 XX
 OS Homo sapiens.
 XX
 PN NO9938972-A2.
 XX
 PD 05-AUG-1999.
 XX
 PF 28-JAN-1999; 99MO-US01619.
 XX
 PR 03-APR-1998; 98US-0080666.
 PR 28-JAN-1998; 98US-0072910.
 PR 24-FEB-1998; 98US-0075954.
 PR 31-MAR-1998; 98US-0080114.
 PR 03-APR-1998; 98US-0080515.
 XX
 PA (CHIR) CHIRON CORP.
 PA (HISE-) HYSEQ INC.
 PI C-Kvenjakov R, Dickson M, Drmanac R, Drmanac S;
 PI Becobedo J, Garcia PD, Garcia V, Glese K, Innis MA;
 PI Jones WL, Kasam A, Kennedy GC, Kita D, Labat J;
 PI Lamson G, Leshkowitz D, Pot D, Randazzo F, Reinhard C;
 PI Stache-Crain B, Sudduth-Klinger J, Williams LT;
 XX
 DR WPI: 1999-494092/41.
 XX
 PT Novel human genes and their expression products which are
 PT differentially expressed in different cell types
 XX
 PS Claim 1; Page 1410-1411; 2479pp; English.
 XX
 CC The present invention describes a library of human polynucleotides
 CC comprising the sequences given in AA21532 to AA21779. Also described is
 CC a method of detecting differentially expressed genes correlated with the
 CC cancerous state of a mammalian cell, comprising detecting at least one
 CC differentially expressed gene product in a test sample from a cell
 CC suspected of being cancerous, where the gene product is encoded by one
 CC of the 5248 polynucleotide sequences given in AA21532 to AA21779. The
 CC polynucleotides can be used as a source of primers and probes, which can
 CC be used for a variety of purposes, e.g. detection of expression levels,
 CC mapping, tissue typing or profiling, forensics, genetic analysis and
 CC detection of polymorphisms. Polypeptides encoded by the polynucleotides
 CC can be used for raising antibodies for experimental, diagnostic and

CC therapeutic purposes. The polynucleotides may also be used to construct
 CC arrays for diagnostics (which may be used to determine function of an
 CC encoded protein); and to detect differences in expression levels between
 CC two cells (e.g. to identify abnormal or diseased tissue in a human, to
 CC identify a genetic predisposition or susceptibility to a disease such as
 CC cancer). The polynucleotides of the invention are especially used in the
 CC diagnosis, prognosis and management of colorectal cancer, breast cancer,
 CC and lung cancer. The polynucleotides can also be used to screen for
 CC peptide analogues and antagonists.

XX Sequence 715 BP; 179 A; 189 C; 178 G; 157 T; 12 other;

Alignment Scores:

Pred. No.:	1 44e-15	Length:	715
Score:	25.00	Matches:	25
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	9.03%	Indels:	0
DB:	20	Gaps:	0

US-09-896-522-2 (1-277) x AA215454 (1-715)

QY 189 ThrPheValIysPProAlaPheGluGluPheCysIeuProThrybTyraIaaspVal 208
 |||||
 Db 470 ACgTtGtCAAGcTtGcCTTGAAGATtCTGCTTGCcAAAGAAgTATGCTGATGtG 529

QY 209 ILeIleProArgGly 213
 |||||
 Db 530 ATCATCCCTAGAGGT 544

RESULT 15

AAFI5785
 ID AAFI5785 standard; cDNA; 1310 BP.

XX AAFI5785;

DT 13-MAR-2001 (first entry)

XX Human prostate cancer antigen nucleotide sequence SEQ ID NO:220.

XX Human; prostate cancer; prostate cancer antigen; detection; diagnosis;
 KW neuroprotective; cytosolic; cardiovascular; immunomodulatory; muscular;
 KW vulnary; gastrointestinal; nephrotropic; antiinfective; gynaecological;
 KW antibacterial; gene therapy; neural; immune; reproductive; renal;
 KW gastrointestinal; pulmonary; cardiovascular; proliferative disorder;
 KW wound; infectious disease; ss.

XX Homo sapiens.

XX WO200055174-A1.

XX 21-SEP-2000.

XX 08-MAR-2000; 2000MO-US05988.

XX 12-MAR-1999; 99US-0124270.

XX (HUMA-) HUMAN GENOME SCI INC.
 PA (ROSE/) ROSEN C A.

XX Rosen CA, Ruben SM;

XX WPI; 2000-587513/55.

XX P-PSDB; AAB56582.

XX Prostate cancer associated gene sequences, referred to as prostate
 PT cancer antigens, useful for treatment, prevention, and diagnosis of
 PT disorders such as prostate cancer -

XX Claim 1; Page 752; 2338pp; English.

XX AAFI5566 to AAFI6505 encode the human prostate cancer associated
 CC proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.

CC The prostate cancer antigens can have neuroprotective, cytosolic,
 CC cardiovascular, immunomodulatory, muscular, vulnary, gastrointestinal,
 CC nephrotropic, antiinfective, gynaecological and antibacterial activities,
 CC and can be used in gene therapy. The prostate cancer antigen
 CC polynucleotides may be used for detection of prostate cancer, chromosome
 CC identification, as chromosome markers, and for numerous other diagnostic
 CC or research purposes. The prostate cancer antigens may be used to treat
 CC disorders such as neural, immune, muscular, reproductive,
 CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
 CC disorders, wounds, and infectious diseases. AAFI5506 to AAFI6514 to
 CC AAB57303 represent sequences used in the exemplification of the present
 CC invention.

XX Sequence 1310 BP; 338 A; 350 C; 352 G; 264 T; 6 other;

Alignment Scores:	2.54e-15	Length:	1310
Pred. No.:	25.00	Matches:	25
Score:	100.00%	Conservative:	0
Percent Similarity:	100.00%	Mismatches:	0
Best Local Similarity:	100.00%	Indels:	0
Query Match:	9.03%	Gaps:	0
DB:	21		

US-09-896-522-2 (1-277) x AAFI5785 (1-1310)

QY 189 ThrPheValIysPProAlaPheGluGluPheCysIeuProThrybTyraIaaspVal 208
 |||||
 Db 789 ACgTtGtCAAGcTtGcCTTGAAGATtCTGCTTGCcAAAGAAgTATGCTGATGtG 848

QY 209 ILeIleProArgGly 213
 |||||
 Db 849 ATCATCCCTAGAGGT 863

Search completed: November 25, 2003, 10:20:43
 Job time : 349 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: November 25, 2003, 09:29:29 ; Search time 3972 Seconds

(without alignments)
2852.962 Million cell updates/sec

Title: US-09-896-522-2
Perfect score: 277
Sequence: 1 MASAGEDCESPAPEADRP.....HPGMLTSGKSHLESSRPH 277

Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2888711 seqs, 20454813386 residues

Word size: 15

Total number of hits satisfying chosen parameters: 50

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODE=frame.p2n.model -DEV=xlh
-Q=/cgn2.1/USPTO.spool/US09896522/runat_21112003_184459_5866/app_query.faeta_1.455
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n_ol1.rge -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bases -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=quality -THR MIN=15 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pro -NOR=ext -HEAR=ext -MINLEN=0 -MAXLEN=200000000
-USER=US09896522@cgn_1.1.3508@runat_21112003_184459_5866 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : GenEmbl :
1: gb_da :
2: gb_htg :
3: gb_in :
4: gb_om :
5: gb_ov :
6: gb_dat :
7: gb_ph :
8: gb_pl :
9: gb_pr :
10: gb_ro :
11: gb_sts :
12: gb_sy :
13: gb_un :
14: gb_vi :
15: em_da :
16: em_fun :
17: em_hum :
18: em_in :
19: em_mu :
20: em_om :
21: em_or :
22: em_ov :
23: em_pac :
24: em_ph :
25: em_pl :
26: em_ro :
27: em_sts :
28: em_un :

29: em_vl :
30: em_hcg_hum :
31: em_hcg_hum :
32: em_hcg_other :
33: em_hcg_mus :
34: em_hcg_pin :
35: em_hcg_rod :
36: em_hcg_mam :
37: em_hcg_vrt :
38: em_sy :
39: em_hcgo_hum :
40: em_hcgo_mus :
41: em_hcgo_other :

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	277	100.0	834	6	AX449219	AX449219 Sequence
2	277	100.0	834	9	AF254133	AF254133 Homo sapi
3	277	100.0	1022	9	AF237290	AF237290 Homo sapi
4	277	100.0	1624	6	AX449217	AX449217 Sequence
5	277	100.0	2160	6	BD157613	BD157613 Primer fo
6	277	100.0	2160	9	AK022317	AK022317 Homo sapi
7	191	69.0	753	6	BD146824	BD146824 Primer fo
8	189	68.2	1395	9	AF125106	AF125106 Homo sapi
9	169	61.0	2072	9	BC015547	BC015547 Homo sapi
10	165	59.6	2228	9	AK057848	AK057848 Homo sapi
11	112	40.4	734	6	AX540411	AX540411 Sequence
12	60	21.7	147492	9	AL358781	AL358781 Human DNA
13	45	16.2	1810	10	MUSURKI	L31783 Mus musculu
14	45	16.2	1959	10	BC025146	BC025146 Mus muscu
15	45	16.2	254993	2	AC078885	AC078885 Mus muscu
16	30	10.8	841	11	BV017303	BV017303 S212P6359
17	30	10.8	952	10	MUSURKIA	L31784 Mus musculu
18	30	10.8	192336	2	AC118474	AC118474 Mus muscu
19	30	10.8	211075	10	AC076974	AC076974 Mus muscu
20	25	9.0	314	10	AB030700	AB030700 Rattus no
21	25	9.0	336	9	BT006860	BT006860 Homo sapi
22	25	9.0	336	12	BT007548	BT007548 Synthetic
23	25	9.0	744	9	AB062451	AB062451 Homo sapi
24	25	9.0	1121	9	AF236637	AF236637 Homo sapi
25	25	9.0	1209	9	BC002906	BC002906 Homo sapi
26	25	9.0	1312	10	BC023789	BC023789 Mus muscu
27	25	9.0	1322	6	AX135546	AX135546 Sequence
28	25	9.0	1354	10	AF236636	AF236636 Mus muscu
29	25	9.0	1784	5	BC045968	BC045968 Danio rer
30	24	8.7	660	11	BV077668	BV077668 S212P6227
31	24	8.7	233210	2	AC097693	AC097693 Rattus no
32	24	8.7	241882	10	ALB08027	ALB08027 Mouse DNA
33	24	8.7	244105	2	AC098897	AC098897 Rattus no
34	24	8.7	254677	2	AC068494	AC068494 Mus muscu
35	19	6.9	383	11	G41891	G41891 SHG-35183
36	17	6.1	655	11	BV077340	BV077340 S212P6009
37	16	5.8	67374	2	AC118933	AC118933 Mus muscu
38	16	5.8	99866	2	AL358115	AL358115 Human DNA
39	16	5.8	120513	2	AC142244	AC142244 Mus muscu
40	16	5.8	148865	2	AL554171	AL554171 Danio rer
41	16	5.8	164246	2	AC016371	AC016371 Homo sapi
42	16	5.8	187700	2	AC131061	AC131061 Mus muscu
43	16	5.8	233137	2	AC127758	AC127758 Rattus no
44	16	5.8	236660	2	AC097934	AC097934 Rattus no
45	15	5.4	831	9	D78335	D78335 Human mRNA

RESULT 1

ALIGNMENTS

AX449219	AX449219	834 bp	DNA	linear	PAT 03-JUL-2002
LOCUS	AX449219	834 bp	DNA	linear	PAT 03-JUL-2002
DEFINITION	Sequence 3 from Patent WO0202761.				
ACCESSION	AX449219				
VERSION	AX449219.1	GI:21697996			
KEYWORDS					
SOURCE					
ORGANISM	Homo sapiens (human)				
REFERENCE	1				
AUTHORS	Glucksmann, M.A.				
JOURNAL	57658, a human uridine kinase and uses thereof				
FEATURES	Patent: WO 0202761-A 3 10-JAN-2002;				
source	Millennium Pharmaceuticals, Inc. (US)				
	Location/Qualifiers				
	1..834				
	/organism="Homo sapiens"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:9606"				
BASE COUNT	194 a 218 c 265 g 157 t				
ORIGIN					
Alignment Scores:					
Pred. No.:	2.6e-298	Length:	834		
Score:	277.00	Matches:	277		
Percent Similarity:	100.00%	Conservative:	0		
Best Local Similarity:	100.00%	Mismatches:	0		
Query Match:	100.00%	Indels:	0		
DB:	6	Gaps:	0		
US-09-896-522-2 (1-277) x AX449219 (1-834)					
QY	1 MetlAsSerAlaGlyGlyGluAspCysGluSerProAlaProGluAlaAspArgProHis	20			
DB	1 ATGGCTTCGGGGGAGCGGAAGACTCCAGAGCCCCGGCGGAGCGCGACCTCGCAC	60			
QY	21 GlnArGpProPheLeuIleGlyValSerGlyGlyThrAlaSerGlyLysSerThrValCys	40			
DB	61 CAGGGGCCCTTCCTGATATAGGGGTAGCGCGGCACATGCGACGGGGAATCCACCGTGT	120			
QY	41 GluLysIleMetCysLeuLeuGlyGlnAsnGluValGluGlnArgGlnArgLysValVal	60			
DB	121 GAGAAAGATCAACGAGGTTCTGTGGACAGAAACGAGTGGAAACAGCGGCACGGAGCGTGTCTC	180			
QY	61 IleLeuSerGlnAspArgPheThrTyrIleValLeuThrAlaGluGlnLysAlaLysAlaLeu	80			
DB	181 ATCTGACCCAGAGACGAGGTCTTCAACAGGCTCTGACGGACAGACAGAGCCAGAGCTTG	240			
QY	81 LysGlyGlnTyrAsnPheAspHisAsProAspAlaPheAspAsnAspLeuMetHisArgThr	100			
DB	241 AAAGGACAGTCAATTTTGACCATCCAGATGCTTTGATATGATTGATGACAGGACT	300			
QY	101 LeuLysAsnIleValGluGlyLysThrValGluValProThrTyrAspPheValThrHis	120			
DB	301 CTGAAGAACATCTGGAGGAGGCAAAACGCTGGAGGTGCCAGACTATGATTTGTGTGACAC	360			
QY	121 SerArgLeuProGluThrThrValValTyrProAlaAspValValLeuPheGluGlyIle	140			
DB	361 TCAAGGTATCCAGAGACCAACGGTGTCTTACCTGTGGACGTGTCTTGTGAGGGCATTC	420			
QY	141 LeuValPheTyrSerGlnGluIleArgAspMetPheHisLeuArgLeuPheValAspThr	160			
DB	421 TTGGTGTCTTACAGCAGAGAGTCCGGGACATGTTCCACTGCGCTCTTGTGTGACACC	480			
QY	161 AspSerAspValArgLeuSerArgArgValLeuArgAspValArgArgGlyArgAspLeu	180			
DB	481 GACTCCGACGTACAGGCTGTCTTCAAGAGATTTCTCCGGGACGTGCGCCGAGGGAGGACCTG	540			
QY	181 GluGlnIleLeuThrGlnTyrThrThrPheValLysProAlaPheGluGluPheGlyLeu	200			
DB	541 GAGCAGATTCTGACCGACAGTACCACTTCTGTGAAGCCGGCTTGAAGAGATTCTGCTG	600			

QY	201	ProthirusylsYrYalaaSPValleleleProRgYlYalaspasmetValalaile	220
Db	601	CCGCAAGAAGATATGCGATGTGATCATCCACGAGAGATGGACATATGTTGCCATC	660
QY	221	AsnleuileValGlnHisileleGlnAspIleleuAnslYlaspIleCyylsYrTPPhIsarG	240
Db	661	AACCTGATCGGCGACACATCCAGGACATTCGTGANTGTGACATCTGCAATATGGCACCGA	720
QY	241	GlYglYseAranglYlAsGSeYrYrYlYArGThrPheserGluProGlyAspHisProGly	260
Db	721	GGAGGGTCCAAATGGCCGAGCTACACCGAACCCTTTCTAGCCAGGGACACACCCCTGGG	780
QY	261	MetLeuTherSerGlylYlYArGSeRHisleuGluSerSerSerArgrProHis	277
Db	781	ATGCTGACCTCTGCGCAACCGTACATTTGGAGTCCAGCAGACAGACCCCGAC	831
RESULT 2			
LOCUS	AF254133	834 bp	mRNA linear PRI 02-MAY-2001
DEFINITION	Homo sapiens uridine kinase mRNA, complete cds.		
ACCESSION	AF254133		
VERSION	AF254133.1	GI:13924749	
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	1 (bases 1 to 834)		
JOURNAL	Ho,Y.S. and Johnson,R.K.		
REFERENCE	Human uridine kinase from prostate cancer cell line (LNCap)		
AUTHORS	Unpublished		
TITLE	2 (bases 1 to 834)		
JOURNAL	Ho,Y.S. and Johnson,R.K.		
AUTHORS	Direct Submission		
TITLE	Submitted (10-APR-2000) Oncology Research, SmithKline Beecham, 709		
JOURNAL	Swedeland Road, King of Prussia, PA 19406, USA		
FEATURES			
source	1..834	Location/Qualifiers	
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	1..834		
	/note="protein kinase"		
	/codon_start=1		
	/product="uridine kinase"		
	/protein_id="AAK49122.1"		
	/db_xref="GI:13924750"		
	/translation="MASAGEDESPAPEDRPHORPFLIGVSGGTASGKSTVCETIM		
	/ELGQNEVEORRQKVVILSDRFKVLTAQKAKLKGQYFHPHPAFDNDLHRLTLK		
	NIVGKTVKEVPLDPTVHSHLPETVVYPPDVVLPEGLVFEQELRDMFHLPLFYDT		
	DSDFRSRLVRIDVRGRDLFOILLTOYTFPKAPFEFLPTKKVADVIIIRGVDMNV		
	AINLIVHIDIDILANGDICMKHRRGSSNRSYKRTFSEGDHPMLTSGKSHLESSRP		
	H"		
BASE COUNT	194 a 219 c 265 g 156 t		
ORIGIN			
Alignment Scores:	2.6e-298	Length:	834
Pred. No.:	277.00	Matches:	277
Score:	100.00%	Conservative:	0
Percent Similarity:	100.00%	Mismatches:	0
Best Local Similarity:	100.00%	Indels:	0
Query Match:	9	Gaps:	0
DB:			
US-09-896-522-2 (1-277) x AF254133 (1-834)			
QY	1	MetAlaserAlaglyglYlUaSPCyglUSeRProAlaProGluAlaAspArGProHis	20
Db	1	ATGCTTCGGGGAGGCGAAGACTGCGAGAGCCCGCGCGGAGCCGACCCGCGAC	60
QY	21	GlnArGProPheLeuIleGlyValSerCylGlyThrAlaSerGlylYseRThrValCyS	40
Db	61	CAGGGCCCTTCCGATAGAGGGGTAGGGCGGCGCATGCCAGCGGAAGTCAGCCGTGTGT	120

Qy	4	GLUylsIleMetGluLeuLeuGluYglInbSngUValIGluInaRGlnAhrgrIyVal	60
Db	121	GAGAAgATCATGGAAGTCTCTGGGACAGAAACAGGTGGAAACAGCGGACGGGAAGGIGTC	180
Qy	61	IleuSerGlnAspArgPheTyrIlySValIleuThrIaGluGlnIlySaIySaIeU	80
Db	181	ATCTGAGCCACGAGACAGGTTCTACAGGTCCTGACGGCAGACAGAAAGGCCAAGGCTTTG	240
Qy	81	LyseGlyGlnTyrAsnPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThr	100
Db	241	AAAGGACAGTCGAATTTGGATCCATCCAAATGCTTTGACAAATGATTTGATGCAGGACT	300
Qy	101	LeuIlySaSnIleValIGluGlyLyseThrValIGluValProThrTyrAspPheValThrHis	120
Db	301	CTGAAGAACATCTGTGAGGCGCAAAACGGTGGAGGTGCCGACCTTAAGATTTTGTACACAC	360
Qy	121	SerArgIleuProGluThrThrValValTyrProAlaAspValValIleuPheGluIlyIle	140
Db	361	TCAAGGTTACAGAGACACCGGTGGTCTACCTCGGAGCGTGGTCTGTGTTGAGGGCATC	420
Qy	141	LeuValPheTyrSerGlnGluIleArgAspMetPheHisIleuArgIleuPheValAspThr	160
Db	421	TTGGTGTTCTACAGCAGAGAGATCCGGGACAGATTCACCTGGCGCTCTCTGTGACACC	480
Qy	161	AspSerAspValArgIleuSerArgArgValIleuArgAspValArgArgGlyYArgAspIleu	180
Db	481	GACTCCGACGTCAGGCTGTCTCGAAGAGTTCCTCGGAGCGTGGCCCGGAGGAGGACCTG	540
Qy	181	GluGlnIleuLeuThrGlnTyrThrThrPheValIlySProAlaPheGluGluIlyuPheCysIleu	200
Db	541	GAGCAGATTTGACGCGAGTACACCACTTCGTAAGCCGGCTCTGAGAGATTTGCGCTG	600
Qy	201	ProThrIlySlyTyrAlaAspValIleIleProArgGlyValAspAsnMetValAlaIle	220
Db	601	CCGACAAAGAAAGTACCGCATGTGATCATCCACAGAGAGTGACAAATATGTTTCCATC	660
Qy	221	AsnIleuIleValGlnHisIleGlnAspIleIleuAsnGlyAspIleCysIlySerPHisArg	240
Db	661	AACTGATCATGTCAGACACATCCAGACATTCGTAATGTGTGATCTTCMAATGGACCGA	720
Qy	241	GlyIySerAsnGlyYArgSerTyrIlyArgThrPheSerGluProGlyAspHisProGly	260
Db	721	GGAGGGTCCATAGGGCGAGGCTACAAAGCGGACCTTTCTGAGCCAGGGGACCAACCTGG	780
Qy	261	MetIeuThrSerGlyIlyYArgSerHisIleuGluSerSerSerArgProHis	277
Db	781	ATGCTGACCTCTGGCAACGGTCACATTTGGAGTCCACACAGACACCCAC	831
RESULT 3			
LOCUS	AF237290	1022 bp	mRNA linear PRI 02-AUG-2001
DEFINITION	Homo sapiens uridine-cytidine kinase 1 (UCK1)	mRNA, complete cds.	
ACCESSION	AF237290		
KEYWORDS	AF237290.1 GI:13506764		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	1 (bases 1 to 1022)		
	Van Rompay,A.R., Norda,A., Lindem,K., Norda,A., Zhu,C., Zheng,X.,		
	Phosphorylation of uridine and cytidine nucleoside analogs by two		
	human uridine-cytidine kinases		
	Mol. Pharmacol. 59 (5), 1181-1186 (2001)		
JOURNAL	21203813		
MEDLINE	11306702		
PUBMED	2 (bases 1 to 1022)		
REFERENCE	Van Rompay,A.R., Lindem,K., Norda,A., Zhu,C., Zheng,X.,		
AUTHORS	Johansson,M. and Karlsson,A.		
	Human uridine-cytidine kinase 1 and 2 : rate limiting enzymes		
TITLE	Unpublished		
JOURNAL	3 (bases 1 to 1022)		
REFERENCE			

AUTHORS	Van Rompay,A.R., Linden,X. Norda,A., Zhu,C., Zheng,X., Johansson,M. and Karlsson,A.
TITLE	Direct Submission
JOURNAL	Submitted (19-FEB-2000) IMPT, Clinical Virology, Huddinge University Hospital, Stockholm 14186, Sweden
FEATURES	Location/Qualifiers
SOURCE	1. .1022
gene	/organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606"
CDS	1. .1022 /gene="UCLK1" 5. .838 /gene="UCLK1" /function="phosphorylates uridine and cytidine to uridine monophosphate and cytidine monophosphate" /codon_start=1 /product="uridine-cytidine kinase 1" /protein_id="AAK28124.1" /db_xref="GI:13506765" /translation="MASAGGEDSCPAEADRPHPRLGVSGGTASGKSTVCEKIMELGQNEVEQRORQVILSDRFYKVLTAQKALAGVNFDPDAFNDLMHRTLKNIVGEKVEVPTDVFTHSRLEPTTVVPAADVLFEEILVFSQEIMDLRLFDVMDSDVRSRVLRLVDVRRRDLBQILTYTTTFVKPAFESECLPTKKYADVILPRGVNWMALNIVGHIDILNGDICKMKHGGSNRSYKRTFSBFDHPGMLTSGKSHLESSNRH"
BASE COUNT	227 a 282 c 324 g 189 t
ORIGIN	
Alignment Scores:	
Pred. No.:	3.16e-298 Length: 1022
Score:	277.00 Matches: 277
Percent Similarity:	100.00% Conservative: 0
Best Local Similarity:	100.00% Mismatches: 0
Query Match:	100.00% Indels: 0
DB:	9 Gaps: 0
US-09-896-522-2 (1-277) x AF237290 (1-1022)	
Qy	1 MetAlaSerAlaGlyGlyGluAspCysGluSerProAlaProGluAlaAspArgProHis 20
Db	5 ATGGCTTCGGGGAGGCGGAAGACTGCGAGAGCCCGCGCGGAGCCGACCTCCGAC 64
Qy	21 GlnArgProPheLeuIleGlyValaSerGlyGlyThrAlaSerGlyLysSerThrValCys 40
Db	65 CAGGGCCCTTCCTCGATAGGGGTAGCGGGCGGCGACTGCCAGCGGGAAGTCACCTGTGT 124
Qy	41 GlnLysIleMetGluLeuLeuGlyGlnAsnGluValGluGlnArgGlnArgLysValaVal 60
Db	125 GAGAAAGATCATGGAGTGTCTGGGACAAGACAGATGTGAACACGCGGACGGAAAGTGTCT 184
Qy	61 IleLeuSerGlnAspArgPheTyrLysValIleuThrLysGluGlnLysAlaLysAlaIleu 80
Db	185 ATCTGTAGCCAGGACAGAGGTTCTACAAAGTCTCTGACGGCAGACAGAGCCAAAGCCTTG 244
Qy	81 LysGlyGlnTyrAsnPheAspHisLeProAspAlaPheAspAsnAspLeuMetHisArgThr 100
Db	245 AAAGACAGTACATATTGTACACATCCAGATGCGCTTGTATATGATTGATGTCACAGAGCT 304
Qy	101 LeuLysAsnIleValaGluGlyLysThrValGluValProThrTyrAspPheValThrHis 120
Db	305 CTGAAGAAACATCGGGAGGCGCAAAACCGTGGAGGTGCGACCTATGATTTGTGACACAC 364
Qy	121 SerArgLeuProGluThrThrValaValTyrProAlaAspValValaLeuPheGluGlyTyr 140
Db	365 TCAAGGTTTACACAGACACACGGTGTCTTACCTCGGACCGTGTCTGTTTGAAGGCACTC 424
Qy	141 LeuValPheTyrSerGlnGluIleArgAspMetPheHisLeuArgLeuPheValaAspThr 160
Db	425 TTGGTGTTCTTACAGCCAGAGATCCGGGACATGTTCACTCGGCCCTTCTGTGTGACACCC 484
Qy	161 AspSerAspValaArgLeuSerArgArgValaLeuArgAspValaArgArgGlyArgAspLeu 180

```

Db      485 GACTCCGACGTCAGGCTGTCTCGAAGAGTTCTCCGGGACGTCGCCGAGGAGGAGACTG 544
Qy      181 GluGlnIleLeuThrGlnTyrThrThrPheValIysProAlaPheGluGluPheCysLeu 200
Db      545 GAGGAGATTCTGACGACGACGACACACACTTCGTGAAGCCGCGCTTCGAGGAGTTCTGCTG 604
Qy      201 ProThrIysIysTyrAlaAspValIleIleProArgIlyValaAspAsnMetValAlaIle 220
Db      605 CCGCAAGAAGAGTATGCGCATGATCATCCCGGAGAGTGGACAATATGTTGCTATC 664
Qy      221 AsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysIysTyrPheArg 240
Db      665 AACCTGATCTGTCGACGACATCTCAGACATTTGATGATGATGATGATGATGATGATGATG 724
Qy      241 GlyIysSerAsnGlyIysSerTyrIlyAspThrPheSerGluProGlyIysAspHisProGly 260
Db      725 GAGAGGTCCTGATGCGGAGGAGCTCAAGCAGGACCTTTCTGAGCCAGGGGACCACTGGG 784
Qy      261 MetLeuThrSerGlyIysArgSerHisLeuGluSerSerArgProHis 277
Db      785 ATGCTGACCTCTGGCAACGGTCACATTTGAGTTCAGCAGCAGACACCCAC 835

RESULT 4
AX449217 1624 bp DNA linear PAT 03-JUL-2002
LOCUS     Sequence 1 from Patent WO0202761.
ACCESSION AX449217
VERSION   AX449217.1 GI:21697994
KEYWORDS
SOURCE    Homo sapiens (human)
ORGANISM  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE
1 Glucksmann, M.A.
  TITLE   Patent: WO 0202761-A 1 10-JAN-2002;
  JOURNAL Millennium Pharmaceuticals, Inc. (US)
          Location/Qualifiers
            1..1624
              /organism="Homo sapiens"
              /mol_type="genomic DNA"
              /db_xref="taxon:9606"
              94..927
                /note="unnamed protein product"
                /codon_start=1
                /protein_id="CAD37771.1"
                /translation="MASAGGEDCSPEADRPHPORFLIGSGTASGSKSTCEKIM
                ELGQNEVEORQRKVVLISODRPFYKULTAQKAKALKQYNPDHDPDFNDLHRTLK
                NIVEGKVEVPTIDFTHSRLPETTVYPADVLFEGILVYSOEIIDMFHLRLFDVT
                DSDVLSRVLADRGRLDQLITQYTTVKKAFEEFCLPTKKYAAVILPREVDNNV
                AINLIYOHIDILNGDICMKWHRGGSNRSYKRTFSEPDHGMILTSGRSHLESSRP
                H"
          CDS
            354 a 427 c 506 g 337 t
              ORIGIN
                BASE COUNT 354 a 427 c 506 g 337 t
                ALIGNMENT Scores:
                  Pred. No.: 4,92e-298 Length: 1624
                  Score: 277.00 Matches: 277
                  Percent Similarity: 100.00% Conservative: 0
                  Best Local Similarity: 100.00% Mismatches: 0
                  Query Match: 100.00% Indels: 0
                  DB: 6 Gaps: 0

US-09-896-522-2 (1-277) x AX449217 (1-1624)

Qy      1 MetIAspSerAlaGlyIyGlyIyAspCysGluSerProAlaProGluAlaAspArgProHis 20
Db      94 ATGGCTTCGGGGAGGAGGAGACTGCGAGAGCCCGCGCGAGGCGGACCGTCCGAC 153
Qy      21 GlnArgProPheLeuIleGlyValSerGlyIyIlyThrAlaSerGlyIysSerThrValCys 40

Db      154 CAGCGGCGCTCTCTGATTAAGGGGTGAGCGGGCGGACGTCGCCAGCGGAGAGTGCACCTGTGT 213
Qy      41 GluIysIleMetGluLeuLeuGlyGlnAsnGluValGluGlnArgGlnArgIysValIle 60
Db      214 GAGAAATGATCATGAGATGTTGTCGGGACAGAAAGAGTGAACAGCGGACGCGGAGGAGTGC 273
Qy      61 IleLeuSerGlnAspArgPheTyrIysValIleuThrAlaGluGlnIlyValAlaValAla 80
Db      274 ATCTGAGCGGAGGACAGGTTCTTAAGAGTCTGACGCGAGAGCAGAAAGGCGCAAGGCTTG 333
Qy      81 LysGlyGlnTyrAsnPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThr 100
Db      334 AAGGACAGTACATTTTATGACATCCAGATGCTTTGATATGATTTGATGACAGAGACT 393
Qy      101 LeuIysAsnIleValGluGlyIysThrValGluValProThrTyrAspPheValThrHis 120
Db      394 CTGAAACATCTCTGAGGAGGCAAAAGCTGAGAGTCCACCATATATTTTGGACACAC 453
Qy      121 SerArgLeuProGluThrThrValIlyTyrProAlaAspValIleuPheGluGlyIle 140
Db      454 TCAAGGTTCACAGAGACACGAGTGTCTACCTGCGGAGGTGTTCTGTTGAGGCGATC 513
Qy      141 LeuValPheTyrSerGlnGluIleArgAspMetPheHisLeuArgLeuPheValAspThr 160
Db      514 TTGGTGTTCACAGCAGGAGGATCCGGGACATGTTCCACTGCGCTTCGTGGACAC 573
Qy      161 AspSerAspValArgLeuSerArgValIleuArgAspValArgArgGlyIysArgAspLeu 180
Db      574 GACTCCGACGTCAGGCTGTCTCGAAGAGTTCCTCGGAGAGTGGCCGAGGAGGAGCTCG 633
Qy      181 GluGlnIleLeuThrGlnTyrThrThrPheValIysProAlaPheGluGluPheCysLeu 200
Db      634 GACCAAGATTCGACGAGACGACACACCTTCGTGAAACCGGCTTCAGAGAGTTCTGCTG 653
Qy      201 ProThrIysIysTyrAlaAspValIleIleProArgGlyValaAspAsnMetValAlaIle 220
Db      694 CCGCAAGAAGAGTATGCGCATGATGATCATCCCGAGAGTGGACATATGTTGCTCATC 753
Qy      221 AsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyIysAspIleCysIysTyrPheArg 240
Db      754 AACCTGATCTGTCAGACATCCAGGACATTCGAAATGTGACATTCGCAAAATGGACCGA 813
Qy      241 GlyIysSerAsnGlyIysSerTyrIlyAspThrPheSerGluProGlyIysAspHisProGly 260
Db      814 GAGGAGTCCATGCGGAGGAGCTTCAAGCGGACCTTTCTGAGCCAGGAGGACCCCTGG 873
Qy      261 MetLeuThrSerGlyIysArgSerHisLeuGluSerSerArgProHis 277
Db      874 ATGCTGACCTCTGGCAACGGTCACATTTGAGTTCAGCAGCAGACACCCAC 924

RESULT 5
BD157613 2160 bp DNA linear PAT 17-JUN-2003
LOCUS     Primer for synthesizing full-length cDNA and use thereof.
ACCESSION BD157613
VERSION   BD157613.1 GI:27863371
KEYWORDS  JP 2002191363-A/12456.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
          1 (base 1 to 2160)
REFERENCE 1 Oca, T., Isogai, T., Nishikawa, T., Hayashi, K., Saito, K., Yamamoto, J.,
          Ishii, S., Sugiyama, T., Wakamatsu, A., Nagai, K., and Otsuki, T.
          Primer for synthesizing full-length cDNA and use thereof
          Patent: JP 2002191363-A 12456 09-JUL-2002;
          HELIX RESPARCH INSTITUTE
          OS Homo sapiens (human)
          PN JP 2002191363-A/12456
          PD 09-JUL-2002
          PF 28-JUL-2000 JP 200280990
          PI TOSHIO OCA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU
          SAITO,

```

PI JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,
 PI KEIICHI NAGAI, TETSUJI OTSUKI
 PC
 C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC
 10,
 PC C12P21/02, C12Q1/68//C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC
 Primer for synthesizing full-length cDNA and use thereof FH Key
 Location/Qualifiers
 (95) .(925) .
 FT CDS
 Location/Qualifiers
 1. .2160
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 BASE COUNT 457 a 591 c 671 g 441 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 6,48e-298 Length: 2160
 Score: 277.00 Matches: 277
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: Gaps: 0
 US-09-896-522-2 (1-277) x BD157613 (1-2160)
 QY 1 MetAlaSerAlaGlyGluValSerGlyLysSerProAlaProGluAlaAspArgProHis 20
 Db 95 ATGGCTTCGGCGGGAGCGGAAGATCGGAGGCCCGCGCGGAGCGCCGCTCCGAC 154
 QY 21 GlnAyrProPheLeuIleGlyValSerGlyGlyThrAlaSerGlyLysSerThrValCys 40
 Db 155 CAGGGCCCTTCCTGATAGGGGTAGGGGGCGGCGCTCCGCGGGAAGTGCAGCGTGT 214
 QY 41 GlnLysIleMetGluLeuGluGlyGlnAsnGluValGluGlnArgGlnArgLysValVal 60
 Db 215 GAGAGATCATGAGGTGCTGGGACAGAACGAGGTGACACGGCGGACGGGAGGTGTCT 274
 QY 61 IleLeuSerGlnAspArgPheTyrrLysValLeuThrAlaGluGlnLysValAlaLys 80
 Db 275 ATCTGTGCGCAGGACAGGCTTCAACAGTCTCTGACGCGCAGGACGAGGCGCAAGCCTTG 334
 QY 81 LysGlyValIleThrAsnPheAspHisProAspAlaPheAspAsnLysMetHisArgThr 100
 Db 335 AAGGACAGTACATTTTGAACATCCAGATGCTTTATATATGATTTGATGACAGGACT 394
 QY 101 LeuLysAsnIleValGluGlyLysThrValGluValProThrTyrrAspPheValThrHis 120
 Db 395 CTGAAAGACATCGTGGAGGGCGAAACGGTGGAGGTGCCGACCTATGATTTTGTGACACAC 454
 QY 121 SerArgLeuProGluThrThrValValTyrrProAlaAspValValLeuPheGluGlyIle 140
 Db 455 TCAAGGTTCACAGACACACGGTGTCTACCTCGGACGCGTGTCTGTTGAGGGCATC 514
 QY 141 LeuValPheTyrrSerGlnGluIleArgAspMetPheHisLeuArgLeuPheValAspThr 160
 Db 515 TTGGTGTTCACAGCCAGGAGATCCGGGACATGTTTCACTGCGCTCTCTGTGAGACACC 574
 QY 161 AspSerAspValArgLeuSerArgArgValLeuArgAspValArgArgGlyArgAspLeu 180
 Db 575 GACTCCGACGTCAAGCGTGTCTCCAAAGAGTCTCCGGGACGTGGCCGACGAGGAGGACCTG 634
 QY 181 GlnGlnIleLeuThrGlnTyrrThrPheValLysProAlaPheGluGluPheCysLeu 200
 Db 635 GAGCAGATTCTGACGACGTACACACCTCTGTGAAGCGGCGCTTGAAGAGTCTGCTGCTG 694
 QY 201 ProThrLysTyrrAlaAspValIleIleProArgLysValAspAsnMetValAlaIle 220
 Db 695 CCGACAAAGAGTATGCGATGATCATCCACGAGAGGTGACAAATATGTTGCCATC 754
 QY 221 AsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysLysTPHISArg 240

Db 755 AACCTGATGCTGACAGACATCCAGACATTCGATGATGTGACATCTGCAAAATGGACCGA 814
 QY 241 GlnGlySerAsnGlyArgSerTyrrLysArgThrPheSerGluProGlyAspHisProGly 260
 Db 815 GAGGGGTCCAAATGGCGGAGCTACAAAGGACCTTTTCTGAGCCAGGAGGACACCCCTGGG 874
 QY 261 MetLeuThrSerGlyLysArgSerHisLeuGluSerSerArgProHis 277
 Db 875 ATCTGACCTCTGCGCAACGCTACATTTGAGATTCACAGACGACGACCCAC 925
 RESULT 6
 AK022317 2160 bp mRNA linear PRI 01-AUG-2002
 LOCUS Homo sapiens cDNA FLJ12255 fis clone MAMMA1001476, highly similar
 DEFINITION to URIDINE KINASE (EC 2.7.1.48).
 ACCESSION AK022317 GI:10433687
 VERSION AK022317.1
 KEYWORDS oligo capping; fis (full insert sequence).
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 BUKARYOTA; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1
 Isogai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y.,
 Nishikawa,T., Nagai,K., Sugano,S., Shiratori,A., Sudo,H.,
 Magatsuna,M., Hosoi,T., Kaku,Y., Kodaira,H., Kondo,H.,
 Sugawara,M., Takahashi,M., Chiba,Y., Ishida,S., Murakawa,K.,
 Ono,Y., Takiguchi,S., Watanabe,S., Kinura,K., Murakami,K.,
 Ishii,S., Kawai,Y., Saito,K., Yamamoto,J., Wakamatsu,A.,
 Nakamura,Y., Negahara,K., Masubo,Y., Ninomiya,K. and Iwayanagi,T.
 NEDO human cDNA sequencing project
 Unpublished
 2 (bases 1 to 2160)
 Isogai,T. and Otsuki,T.
 Direct Submission
 Submitted (23-AUG-2000) Takao Isogai, Helix Research Institute,
 Genomics Laboratory, 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
 (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
 NEDO human cDNA sequencing project supported by Ministry of
 International Trade and Industry of Japan; cDNA full insert
 sequencing; Research Association for Biotechnology; cDNA library
 construction, 5'- & 3'-end one pass sequencing and clone selection;
 Helix Research Institute (supported by Japan Key Technology Center
 etc.) and Department of Virology, Institute of Medical Science,
 University of Tokyo.
 FEATURES
 source
 1. .2160
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MAMMA1001476"
 /issue_type="Mammary gland"
 /clone_lib="MAMMA1"
 /note="Cloning vector: pME186FL3"
 95. 928
 /note="unamed protein product"
 /codon_start=1
 /protein_id="BAB14010.1"
 /db_xref="GI:10433688"
 /translation="MASAGDECPAPADRPORPFLIGVSGGTASGKSTVCEKIM
 ELLGQNEVROKRVYILISODRFYVLTAEOKAKLKGQYNDHPDAPNDMLHRLK
 NIVEGRTVPEPTDPTVTHSRLEPTTVYVADYVLEBGILVPSQELRDPFHRLPDT
 DSDVLRSLVDVRGRDLQILTOYTFVPAPEEFCLPYTKTADVILPQGVDMV
 AINLVQHLQDILNDGICWNRGSGNGRSYKRTFSEPDHPMLTSGKSHLESSRP
 H"
 BASE COUNT 457 a 591 c 671 g 441 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 6,48e-298 Length: 2160
 Score: 277.00 Matches: 277
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0

Db 575 GACTCCGACGTCAGGCTGTCTCGAAGATTCTCCGGGACGTCGCGAGGAGGAGCCTG 634
 Qy 181 GluGlnIleuThrGlnTyrThrThrpheVal 191
 Db 635 GAGCAGATTCTGACGAGTACACCACTTGTGT 667
 RESULT 8
 AF125106 1395 bp mRNA linear PRI 07-FEB-2002
 LOCUS Homo sapiens uridine kinase mRNA, complete cds.
 DEFINITION AF125106
 VERSION AF125106.1 GI:18568108
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 1395)
 AUTHORS Xin,Y.R., Yu,L. and Zhao,S.Y.
 TITLE Cloning of a new human cDNA similar to Mus musculus uridine kinase
 JOURNAL mRNA
 REFERENCE 2 (bases 1 to 1395)
 AUTHORS Ding,J.B., Yu,L. and Zhao,S.Y.
 TITLE Direct Submission
 JOURNAL Submitted (02-FEB-1999) Lab of Human Gene Research, Institute of
 Genetics, Pudan University, No. 220 Handan Rd., Shanghai 200433,
 People's Republic of China
 FEATURES
 source
 1. 1395
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 8. 838
 /codon_start=1
 /product="uridine kinase"
 /protein_id="AAL75943.1"
 /db_xref="GI:18568109"
 /translation="MASAGSEGARARARARPHORPFLIGVSGTASGKSTVEKIME
 LLDQNEVEQRHGVILSDQRFKVLAEKAKALKQYVDPDPAFNDMLHRTLN
 IVEGKTVEPTVDFVTHSLPFTVVPAFVAFEGILVPSQIRDFHLRFVDSD
 SVRLSRVLRVDRGRDLQILQYTFPAFEEFCLPTKYADVILPRGVNMA
 INLVQHIQDILNDICMKHRRGSGNGRTKRTSEPDHDMGLTSGKSHLESSRHH
 "

BASE COUNT 325 a 352 c 411 g 307 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 4.24e-200 Length: 1395
 Score: 189.00 Matches: 258
 Percent Similarity: 98.47% Conservaive: 0
 Best Local Similarity: 98.47% Mismatches: 2
 Query Match: 68.23% Indels: 4
 Gaps: 0
 US-09-896-522-2 (1-277) x AF125106 (1-1395)
 Qy 18 ArgProHisGlnArgProPheLeuIleGlyValSerGlyGlyThrAlaSerGlyLysSer 37
 Db 56 CGTCCGACACGAGCGTCCCTCTCTATAGGGGTGACGCGGCACTGCCAGCGGGAAGTCG 115
 Qy 38 ThrValGlyGlyValIleMetGluLeuLeuGlyGlnGlnGluValGluGlnArg 57
 Db 116 ACCGTGTGTGAAGATCATGAGTCTCTGGGACAGACAGAGGTGGAACAGCGGCA-CGG 174
 Qy 58 -LysValValIleLeuSerGlnAspArgPheTyrLysValIleuThrAlaGlnGlnLysAl 77
 Db 175 TAGAGTGTATCTCTGACGACAGCAAGTTCTTACAAGTCTTACGCGACAGCAAGAGGCC 234
 Qy 77 alyValAlaLeuLysGlyGlnTyrAsnPheAspHisProAspAlaPheAspAsnAspLeu 97
 Db 235 CAAGGCGCTTGAAGAGACAGTAAATTGTCATCCAGATGCCCTTGATATGATTGAT 294

Qy 97 cHisArgThrLeuLysAsnIleValGlnGlyLysThrValGluValProThrTyrAsp 117
 Db 295 GCACAGACTCTTAAGAAACATCTGTGAGGGCAAAACGTGAGAGTGTCCACCTATATTT 354
 Qy 117 eValIrrHisSerArgLeuProGluThrThrValValIlyrProAlaAspValIleuPh 137
 Db 355 TGTACACACTCAAGGTACACAGAGCAACGGGTGTCTACCTCGGAGCGTGTCTGTT 414
 Qy 137 eGluGlyIleLeuValPheTyrSerGlnGluIleAlaGAspMetPheHisLeuArgLeuPh 157
 Db 415 TGAAGGACATCTTGTTGTCTTACAGCCAGAGAGATCCGGACATGTTCCACTGGCCCTTT 474
 Qy 157 eValAspThrAspSerAspValAlaArgLeuSerArgValIleuArgAspValArgArg 177
 Db 475 CGTGCACACCGACTCCGACGTCAGGCTGTCTCAAGATTTCTCCGGACGTCGGCGCGG 534
 Qy 177 yArgAspLeuGluGlnIleLeuThrGlnTyrThrThrPheValLysProAlaPheGlu 197
 Db 535 GAGGAGACTGGACAGATCTTGAACGATACACCACTTCGTGAACCGGCTTCGAGGA 594
 Qy 197 upheCyLeuProThrLysLysTyrAlaAspValIlelleProArgGlyValAspAsnMe 217
 Db 595 GTTCTGCTGCGCCAGCAAGAGATATGCCATGTGATCATCCCAACGAGGTGACAAATAT 654
 Qy 217 tValAlaIleAsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysLy 237
 Db 655 GGTTCGATCATCCTTATCTGACGACATCCGACATTTCTGAAAGTGATCTGCAAT 714
 Qy 237 sTrpHisArgGlyGlySerAsnGlyArgSer-TyrLysArgThrPheSerGluProGly 257
 Db 715 ATGGACCCGAGGAGGGGTCCATGAGGGGAC-GRACAAGGGAGACTTTTTCGACCAAGGGG 773
 Qy 257 sPHisProGlyMetLeuThrSerGlyLysArgSerHisLeuGluSerSerArgProH 277
 Db 774 ACCACCTGGAGATGATGATCTTGGCAACAGGTACATTGAGTCCAGACGACAGACCCC 833
 Qy 277 ts 277
 Db 834 AC 835

RESULT 9
 BC015547 2072 bp mRNA linear PRI 04-OCT-2001
 LOCUS Homo sapiens, similar to uridine-cytidine kinase 1, clone MGC:9668
 DEFINITION IMAGE:3845821, mRNA, complete cds.
 ACCESSION BC015547
 VERSION BC015547.1 GI:15930229
 KEYWORDS MGC.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 2072)
 AUTHORS Strausberg,R.
 TITLE Direct Submission
 JOURNAL Submitted (01-OCT-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgabs-r@mail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome
 Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: <http://www.sbgc.stanford.edu>
 Contact: (Dickson, Mark) mcd@paxli.stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
 R. M.

Clone distribution: MGC clone distribution information can be found

through the I.M.A.G.E. Consortium/JMLN, at: <http://image.llnl.gov>
 Series: IRAX plate: 20 Row: f Column: 5
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 13899252.

FEATURES

Location/Qualifiers

1..2072

CDS

/organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:9668 IMAGE:3845821"
 /tissue_type="Colon, adenocarcinoma"
 /clone_id="NH_MGC_65"
 /lab_host="DH10B"
 /note="Vector: pCMW-SPORT6"
 87..692
 /codon_start=1
 /product="similar to uridine-cytidine kinase 1"
 /protein_id="AAH15547.1"
 /db_xref="GI:15930230"
 /translation="MASAGEDCESSPAEADRPQRPFLIGSGTASGKSTVCEKIM
 ELIGNEVEOROKRVILISODRFYKVLAEQKAKLQYFDPDPAFDNDLMHRTLK
 NIVEKTVETPTDVTHTSRLLPETHVYPAVULPEGLTPYSGEIRDMFLRLPVT
 DSDVRLSRKDEKRCDFPTRSQGYGCHQDPRAAHFGHSEW"

BASE COUNT 455 a 559 c 637 g 421 t
 ORIGIN

Alignment Scores:

Pred. No.: 1..166-177 Length: 2072
 Score: 169.00 Matches: 169
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 61.01% Indels: 0
 DB: 9 Gaps: 0

US-09-896-522-2 (1-277) x BC015547 (1-2072)

QY 1 MetAlaSerAlaGlyGlyAlaSPcGsluSerProAlaProGluAlaAspArgProHis 20
 Db 87 ATGGCTTCGGCGGGAGCGGAAGACTCGAGAGCCCGCGCGGAGCGGACCGTCCGAC 146
 QY 21 GlnArgProPheLeuIleGlyValSerGlyThrAlaSerGlyLysSerThrValCys 40
 Db 147 CAGCGGCCCTTCCTGATAGGGGTGAGCGCGGCACCTGCACGGGAATCGACCGTGTGT 206
 QY 41 GluLysIleMetGluLeuLeuGlyGlnAsnGluValGluGlnArgGlnArgLysVal 60
 Db 207 GAGAAAGATCATGAGATTCTCGGACAGAACAGAGTGGAAACCGCGACGGAAAGTGTGTC 266
 QY 61 IleLeuSerGlnAspArgPheTyrIleValLeuThrAlaGluGlnLysAlaLysAlaLeu 80
 Db 267 ATCTGAGCCAGAGCAGAGTTCACAAAGTCTCTGACGGCAGAGCAGAAAGCCAGGCTTG 326
 QY 81 LysGlyGlnTyrAsnPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThr 100
 Db 327 AAAGACAGTCAATTGACCATCCAGATGCTTTGATTAATGATTGATTCACAGAGACT 386
 QY 101 LeuLysAsnIleValGluGlyLysThrValGluValProThrTyrAspPheValThrHis 120
 Db 387 CTGAAGAACATCTCGTGAAGGCGAAACGGTGAAGGTGCCGACCTTATGATTTTGTACACAC 446
 QY 121 SerArgLeuProGluThrThrValValTyrProAlaAspValValLeuPheGluGlyIle 140
 Db 447 TCAAGGTACCAAGAGACCAAGGTGGTCTACCTCGGAGCGTGGTCTTTGAGGGATTC 506
 QY 141 LeuValPheTyrSerGlnGluIleArgAspMetPheHisLeuValGluLeuPheValAspThr 160
 Db 507 TTGGTGTCTACAGCCAGAGAGATCCGGGACATGTTTCCACTGCGCTCTTGTGTGACACC 566
 QY 161 AspSerAspValArgLeuSerArgArg 169
 Db 567 GACTCCAGCTCAGGCTGTCTGAGAGA 593
 RESULT 10

AK057848

LOCUS

DEFINITION

AK057848 2228 bp mRNA linear PRI 31-OCT-2001

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AK057848 2228 bp mRNA linear PRI 31-OCT-2001
 Homo sapiens cDNA FLJ25119 fis, clone CBR05878, highly similar to
 URIDINE KINASE (EC 2.7.1.48).
 AK057848
 AK057848.1 GI:16553809
 oligo capping; fis (full insert sequence).
 Homo sapiens (human)
 Homo sapiens
 Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1
 Tashiro, H., Yamazaki, M., Watanabe, K., Kumagai, A., Itakura, S.,
 Fukuzumi, Y., Fujimori, Y., Komiyama, M., Suzuki, Y., Hata, H.,
 Nakagawa, K., Mizuno, S., Morinaga, M., Kawamura, M., Sugiyama, T.,
 Irie, R., Otsuki, T., Sato, H., Nishikawa, T., Sugiyama, A.,
 Kawakami, B., Nagai, K., Isegai, T. and Sugano, S.
 NEDO human cDNA sequencing project
 Unpublished
 2 (bases 1 to 2228)
 Sugano, S. and Suzuki, Y.
 Direct Submission
 Submitted (24-OCT-2001) Sumio Sugano, Institute of Medical Science,
 University of Tokyo, Laboratory of Genome Structure, Human Genome
 Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan
 (E-mail:cdna@ims.u-tokyo.ac.jp; Tel:81-3-5449-5286,
 Fax:81-3-5449-5416)
 NEDO human cDNA sequencing project supported by Ministry of
 Economy, Trade and Industry of Japan; cDNA full insert sequencing:
 Research Association for Biotechnology (RAB); cDNA library
 construction and 5'-end one pass sequencing: Institute of Medical
 Science, University of Tokyo, Laboratory of Genome Structure, Human
 Genome Center; 3'-end one pass sequencing: RAB; clone selection for
 full insert sequencing: RAB and Helix Research Institute.
 Location/Qualifiers

FEATURES

source

1..2228

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CBR05878"

/tissue_type="brain"

/clone_id="CBR"

/note="cloning vector: pME18SFL3"

BASE COUNT

478 a 598 c 696 g 456 t

ORIGIN

Alignment Scores:

Pred. No.: 3..546-173 Length: 2228
 Score: 165.00 Matches: 165
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 59.57% Indels: 0
 DB: 9 Gaps: 0

US-09-896-522-2 (1-277) x AK057848 (1-2228)

QY 37 SerThrValCysGluLysIleMetGluLeuLeuGlyGlnAsnGluValGluGlnArgGln 56
 Db 214 TCACCGTGTGTAGAAAGATCATGAGTTCCTGGACAGAACAGAGTGGAAACCGCGCAG 273
 QY 57 ArgLysValValIleLeuSerGlnAspArgPheTyrLysValLeuThrAlaGluGlnLys 76
 Db 274 CGAAGGTGTCTATCTCGAGCCAGAGCAGAGTCTTCAAGGTCTCGACGCGAGCAGAG 333
 QY 77 AlaLysAlaLeuLysGlyGlnTyrAsnPheAspHisProAspAlaPheAspAsnAspLeu 96
 Db 334 GCCAAGGCCCTTGAAGAGCAGTCAATTGTCACATCCAGATGCCCTTGAATATGATTTC 393
 QY 97 MetHisArgThrLeuLysAsnIleValGluLysThrValGluValProThrTyrAsp 116
 Db 394 ATGCACGAGACTGTGAAGAACATCTGAGAGGCGAAACGGTGAAGTGCAGACTTATGAT 453
 QY 117 PheValThrHisSerArgLeuProGluThrThrValValTyrProAlaAspValValLeu 136

```

Db      454 TTGTGACACACTCAAGTTACAGACAGACCGTGTCTACCTCGGACGTTCTG 513
Qy      137 PhGluGluYlleuValPheTYSerGlnGluIleArgApwecPheH1sleuArgleu 156
Db      514 TTGAGGGGACATCTTGTTGTTCTACAGCCAGAGATCCGGAGACATGTTCCACCTCGCGCTC 573
Qy      157 PheValAspThrAspSerAspValArgLeuSerArgArgValIleuArgAspValArgArg 176
Db      574 TTGTGGACACCGACCTCGACGTCGAGGCTGTCTGAAAGATTCTCCGGAGCGTGGCGGA 633
Qy      177 G1ArgAspLeuGluGlnIleLeuThrGlnTyThrPheValIysProAlaPheGlu 196
Db      634 GGGAGGGACCTGGAGCGAGATTTCAGACCGACATCCACTTCGGAAGCCGCGCTTCGAG 693
Qy      197 GluPheCysLeuPro 201
Db      694 GAGTTCTGCGCTGCGC 708

RESULT 11
AX540411
LOCUS      AX540411      734 bp      DNA      linear      PAT 23-NOV-2002
DEFINITION Sequence 23 from Patent WO02055738.
ACCESSION  AX540411
VERSION     AX540411.1  GI:25273437
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS     Panzer, S.R., Lincoln, S.E., Altue, C.M., Dufour, G.E., Hillman, J.L.,
            Jones, A.L., Dam, T.C., Liu, T.F., Harris, B., Flores, V., Patis, A.,
            Marwaha, R., Chen, A.J., Chang, S.C., Gerstlin, E.H., Peralta, C.H.,
            David, M.H. and Lewis, S.A.
TITLE        Molecules for disease detection and treatment
JOURNAL      Patent: WO 02055738-A 23 18-JUL-2002;
            Incyte Genomics, Inc. (US)
FEATURES
            Location/Qualifiers
            source          1..734
                        /organism="Homo sapiens"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:9606"
                        /note="Incyte ID No: LI:235557.12:2001JUN12"
BASE COUNT  166 a 181 c 226 g 154 t
ORIGIN
Alignment Scores:
Pred. No.:      1 28e-114      length:      734
Score:          112.00      Matches:      112
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    40.43%      Indels:      0
                        Gaps:      0
US-09-896-522-2 (1-277) x AX540411 (1-734)
Qy      90 AspAlaPheAspAsnAspLeuMetHisArgThrIleuIysAsnIleValGluGlyLysThr 109
Db      14 GATGCCCTTGATATGATTGATGACACAGACTCTGAAGAACATCGTGAGGGCAAAACG 73
Qy      110 ValGluValProThrTYrAspPheValThrHisSerArgLeuProGluThrThrValVal 129
Db      74 GTGAGAGTGCAGCACTATGATTTGTGACACACTCAAGATTACCGAGACACGTTGCTC 133
Qy      130 TYrProAlaAspValValIleuPheGluGlyIleuValPheTYSerGlnGluIleArg 149
Db      134 TACCTCGGACGCGGTCTGTTTGAAGGCATCTTGTTGTTTACACCCAGAGATCCGG 153
Qy      150 AspMetPheH1sleuArgleuPheValAspThrAspSerAspValArgleuSerArgArg 169
Db      194 GACATGTTCCACCTGCGCTCTTCGTGAGACACCGACTCCAGCGTCAAGCTGTTCCGAAGA 253

```

```

Qy      170 ValIeuArgAspValArgArgGlyArgAspLeuGluGlnIleLeuThrGlnTYrThr 189
Db      254 GTTCTCCGGAGCGTGGCCGAGAGGAGGACCTGAGACATTTCTGACGATACACACC 313
Qy      190 PheValIysProAlaPheGluGluIleuPheCysLeuPro 201
Db      314 TTGTGAGACCGGCTTCGAGAGATTCTGCTGCTCCG 349

RESULT 12
AL358781/c
LOCUS      AL358781      147492 bp      DNA      linear      PRI 06-OCT-2001
DEFINITION Human DNA sequence from clone RP11-334J6 on chromosome 9, complete
            sequence.
ACCESSION  AL358781
VERSION     AL358781.19  GI:13751418
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
            1 (bases 1 to 147492)
REFERENCE
AUTHORS     Corby, N.
TITLE        Direct Submision
JOURNAL      Submitted (06-OCT-2001) Sanger Centre, Hinxton, Cambridgehire,
            CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
            requests: clonerequest@sanger.ac.uk
COMMENT
            On Apr 21, 2001 this sequence version replaced gi:13396549.
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submision
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest. The following
            abbreviations are used to associate primary accession numbers given
            in the feature table with their source databases: Em: EMBL; Sw:
            SWISSPROT; Tr: TrEMBL; Wp: WORMPEP; Information on the WORMPEP
            database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping
            Group. Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr9
            RP11-334J6 is from the library RPCT-11.2 constructed by the group
            of Pieter de Jong. For further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pBAC3.6
            This sequence is the entire insert of clone RP11-334J6 The true
            left end of clone RP11-40A7 is at 113870 in this sequence. The true
            right end of clone RP11-643B14 is at 63282 in this sequence.

FEATURES
            Location/Qualifiers
            source          1..147492
                        /organism="Homo sapiens"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:9606"
                        /chromosome="9"
                        /clone="RP11-334J6"
                        /clone_11b="RPCT-11.2"
                        /clone_11b="RPCT-11.2"
            misc_feature    30143..30199
                        /note="Sequence from overlapping clone RP11-643B14
                        (AL354855). Assembly confirmed by restriction digest."
BASE COUNT  32992 a 36290 c 37577 g 40633 t
ORIGIN
Alignment Scores:
Pred. No.:      1.68e-54      length:      147492
Score:          60.00      Matches:      60

```

Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	21.66%	Indels:	0
DB:	9	Gaps:	0
US-09-896-522-2 (1-277) x AL358781 (1-147492)			
QY	218	VALAIAIAAsnleuileValGInHsIIeGInAspIIeLeuAnGlyAspIIeCyGlyAs	237
Db	129488	GTTGGCATCAACCTGATCGCGACATCCAGACATTCCTGATGATGATGATCTGC	129429
QY	238	TPHISARGIGLVserAnGlyArgserTYrLYeArGthrpheserGpUProGlyAsp	257
Db	129428	TGGCACCGAGAGGATTCATGGGCGAGCTCAAGGAGCCTTTCTGACCGAGGGAC	129365
QY	258	HisProGlyMeLeuThiSerGlyLYsArgSerHisLeuGluSerSerSerATpProHis	277
Db	129368	CACCCGGATGCTGACCTCTGGCAACAGGTGACATTTGGAGTCCAGACACCCAC	129309
RESULT 13			
MUSURKI		1810 bp	linear
LOCUS			
DEFINITION	Mus musculus uridine kinase mRNA, partial cds.		
ACCESSION	L31783		
VERSION	L31783.1	GI:471980	
KEYWORDS			
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
REFERENCE			
AUTHORS	1 (bases 1 to 1810)		
TITLE	Traut,T.W.		
JOURNAL	The functions and consensus motifs of nine types of peptide		
MEDLINE	segments that form different types of nucleotide-binding sites		
PUBMED	Eur. J. Biochem. 222 (1), 9-19 (1994)		
REFERENCES	8200357		
AUTHORS	2 (bases 1 to 1810)		
TITLE	Ropp,P.A. and Traut,T.W.		
JOURNAL	Cloning and expression of a cDNA encoding uridine kinase from mouse		
MEDLINE	brain		
PUBMED	97108719		
FEATURES	8951040		
source	Location/Qualifiers		
	1..1810		
	/organism="Mus musculus"		
	/mol_type="mRNA"		
	/db_xref="taxon:10090"		
	/tissue_type="brain"		
	<1..783		
	/EC_number="2.7.1.48"		
	/standard_name="uridine-cytidine kinase"		
	/codon_start=1		
	/product="uridine kinase"		
	/protein_id="AA050568.1"		
	/db_xref="GI:471981"		
	/translation="RPPRPFLIGVSGTASGKSTVCEKIMELGONEVDRRORKLVI		
	HSRDFEYKLTAEAKAKALGQVNFPHDPAFNDMLKTLXNIVEKTYEPTVPTV		
	HSRLPETTVVPADVLPEGLIVPTQERIDMFHLRLPDTSDVRLSRVLRDVORG		
	RDLDELITQYIAFVKPAFEFCLPTKIVADVIIPRGVDNVAIINLVQHIDILNSDL		
	CKRHGGPGNRNHKRTFPPGGHPGLATGKSHLESSSRPH"		
	join(37..63,412..426,451..465)		
	/bound_moiety="ATP at catalytic site"		
	1790..1795		
BASE COUNT	411 a 482 c 512 g 405 t		
ORIGIN			
Alignment Scores:			
Pred. No.:	1.25e-39	Length:	1810
Score:	45.00	Matches:	164
Percent Similarity:	95.35%	Conservative:	0
Best Local Similarity:	95.35%	Mismatches:	4

Query Match	16.25%	Indels:	8
DB:	10	Gaps:	0
US-09-896-522-2 (1-277) x MUSURKI (1-1810)			
QY	67	PheTrrLySVaLleuThrAlaGluGlnLySaIalysAlaLeuLySGlyGlnTyrAsn	86
Db	148	TTCTACAAAGGTTCTGCAGCGCCGAGACGAGGCGCAAGGCTTTGAAAGGCGACATACATTTT	207
QY	87	AspHisProAspAlaPheAspAsnAspLeuMetHisArg-ThrLeuLySaenileValG1	106
Db	208	GACCAACCCAGATGCTTTGGATTAAGATGTGATGCACA-GACCTGAAAACATTTGTGGA	266
QY	106	UGLYLySThrValGluValProThrTyrAspPheValThrHisSerArgLeuProGluTh	126
Db	267	AGGCAAACTGTCGAGGTCCTTACCTATGATTTTGTGACCCACATCAAGGTACAGAGAC	326
QY	126	rThrValValTyrProAlaAspValValLeuPheGluGlyIleuValPheTyrSer-G	146
Db	327	CAGCTGGTCTTACCCAGCTGATGAGTGTTCTTGTGAGGCGACTTGGATTCTACAC-CC	385
QY	146	InGluIleArgAspMetPheHisLeuArgLeuPheValAspThrAspSerAspValArgL	166
Db	386	AGGAGATCCGGGACATGTTCCACTGCGCCCTTTGTGACACAGACTGTATGTAGGC	445
QY	166	euserArgArgValleuArgAspValArg-ArgGlyArgAspLeuGluGlnIleuThr	185
Db	446	TGTCGGAAGGTTCTCCGGAGTGTCGACAGAGAGGAGCCTGAGCAGATCTGACT	504
QY	186	GlnTyrThrThr-rheValLySProAlaPheGluGlnPheLySerProThrLyS	205
Db	505	CAGTACACCGC-CTTTGTGAACACAGCCTTTGAGAGAGTTCTGCTCGCGATTAAGAA	563
QY	205	rAlaAspValIleIleProArgGlyValAspAsnMetValAlaIleAsnLeuIleValG1	225
Db	504	CGCTGAGCTGATCATCCTCGAGAGAGTTGATTAATGTGGTGCATCAACCTGATCTGCA	623
QY	225	nhisIleGlnAspIleLeuAsnGlyAsp	224
Db	624	ACACATCCAGAGACATCTCAACGGGGAC	651
RESULT 14			
LOCUS	BC025146	1959 bp	mRNA linear ROD 16-APR-2003
DEFINITION	Mus musculus uridine monophosphate kinase, mRNA (CDNA clone		
ACCESSION	BC025146		
VERSION	BC025146.1	GI:19263563	
KEYWORDS	MG.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
REFERENCE	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
AUTHORS	1 (bases 1 to 1959)		
	Strausberg,R.L., Fellings,E.A., Grouse,L.H., Derge,J.G.,		
	Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,		
	Alschuler,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,		
	Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,		
	Ditchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,		
	Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,		
	Schaefer,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,		
	Carinci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,		
	Abramson,R.D., Mullany,S.J., Bosak,S.A., McEwan,P.J.,		
	McKernan,K.J., Malek,J.A., Gunnarsson,P.H., Richards,S.W.,		
	Wotley,K.C., Hale,S., Garcia,A.M., Gay,L.D., Hulyk,S.W.,		
	Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,		
	Fahney,J., Helton,E., Kettelman,M., Madan,A., Rodriguez,S.,		
	Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,		
	Bouffard,G.C., Blakesley,R.W., Touchman,J.W., Green,E.D.,		
	Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,		
	Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalins,D.E.,		
	Schmeh,A., Schein,J.E., Jones,S.J., and Marz,M.A.		
	Generation and initial analysis of more than 15,000 full-length		

human and mouse cDNA sequences
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 MEDLINE 22388257
 PUBMED 12477932
 REFERENCE 2 (bases 1 to 1959)
 AUTHORS Strausberg, R.
 TITLE Direct Submission
 JOURNAL Submitted (05-MAR-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 COMMENT Contact: MGC help desk
 Email: gcgaps-remail.nih.gov
 Tissue Procurement: Jeffrey E. Green, M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center
 Center code: BCM-HGSC
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
 Contact: amg@bcm.tmc.edu
 Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Louised, H., Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nannavati, A.N., Gibbs, R.A.

FEATURES
 source .
 1..1959
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="MGC:36231 IMAGE:4913412"
 /tissue_type="Salivary gland, 10 week old female mouse"
 /clone_id="NCI CGAP_SG2"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"
 1..1959
 /gene="Umpk"
 /db_xref="LocusID:22245"
 /db_xref="MGI:98904"
 57..890
 /codon_start=1
 /product="Umpk protein"
 /protein_id="AAH25146.1"
 /db_xref="GI:19263564"
 /db_xref="LocusID:22245"
 /translation="MASAGGSGESAPADPPROPFLIGSGTASGKSTCEKIM ELIAGVNRORRKLVILODFKYLTAOKKAKGPNPHPADFNDLHMKITX NIVEGTVREPTDFTVTHSLPRTVVPADVLFEGLVFTYQERIDMHLAFDPT DSVLRSRRLVQGRDLQILITOTATVAKAFEFECPTKYKADAVIIPKGVNNV AINLVQHIQDILNGDLCRRHGGPNRNKRTPPEGDHPGLATGKRSLSSSRP H"

BASE COUNT 466 a 511 c 561 g 421 t
 ORIGIN

Alignment Scores:
 Pred. No.: 1.34e-39 Length: 1959
 Score: 45.00 Matches: 164
 Percent Similarity: 95.35% Conservative: 0
 Best Local Similarity: 95.35% Mismatches: 4
 Query Match: 16.25% Indels: 8
 DB: 10 Gaps: 0

US-09-896-522-2 (1-277) x BC025146 (1-1959)

Qy 67 PhetyrlvsValleuThrrAlaGluInLyAlaLysAlaLeuYsgLyGlnTyraNphe 86

|||||
 Db 255 TTCTACAGAGTTCTGACGGCCGAGAGGCCAAGGCTTTGAAGGAGCATATTTT 314
 Qy 87 AsphIsProAspAlaPheAspAsnAspLeuMetHisArg-ThrLeuLysAsnIleValGI 106
 Db 315 GACCACCCAGATCTTTTGATATGATCTGATCACA-A-GACCTTGAAAAACATTTGTGA 373
 Qy 106 uGlyLvsThrValGluValProThrTyraAspPheValThrHisSerArgLeuProGIuTh 126
 Db 374 AGGCAAACTGTGAGAGTCCCTACCTATGATTTTGACCCACTCAAGGTTTACCAAGAC 433
 Qy 126 rThrValValTyraProAlaAspValValLeuPheGluGlyIleLeuValPheTyrsr-G 146
 Db 434 CACTGTGCTTACCAAGCTGATGTGTTCTGTTCGAGGGAGATTGGTATTTACAC-CC 492
 Qy 146 InGluIleArgAspMetPheHisLeuArgLeuPheValAspThrAspSerAspValArgL 166
 Db 493 AGGAGATCCGGGACATGTTCCACCTCGCCTCTTTTGACACAGACTCTGATGTTAAGGC 552
 Qy 166 euserArgArgValLeuArgAspValArg-ArgGlyArgAspLeuGluGlnIleLeuThr 185
 Db 553 TGTTCTGAAGAGTTCTCCGGATGTGCA-ACGAGGAAGGACCTGAGCAGATCTGACT 611
 Qy 186 GlnTyrThrThr-PheValIysProAlaPheGluGluPheCysLeuProThrIysIsty 205
 Db 612 CAGTACACGCC-CTTTGTGAACACACAGCTTTGAGAGTTCTGCTCCGACCTAAGAAATA 670
 Qy 205 rAlaAspValIleIleProArgGlyValAspAsnMetValAlaIleAsnLeuIleValGI 225
 Db 671 CGCTGACGTGATCATCTCTCGAGAGTGTATATATGTGGCCATCAACCTGATCGTGA 730
 Qy 225 nHisIleGlnAspIleLeuAsnGlyAsp 234
 Db 731 ACAATCCAGACATCTCTCAACGGGAC 758

RESULT 15
 AC078885/c 254993 bp DNA linear HTG 15-MAY-2002
 LOCUS
 DEFINITION Mus musculus chromosome 6 clone RP23-117723 strain C57BL/6/J,
 WORKING DRAFT SEQUENCE, 44 unordered pieces.

ACCESSION
 AC078885 GI:18376842
 VERSION
 HTG; HTGS PHASE1; HTGS DRAFT.
 KEYWORDS
 Mus musculus (house mouse)
 SOURCE
 ORGANISM
 Mus musculus

REFERENCE
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 254993)
 Montgomery, K.T., Grills, G., Han, J., Lee, E., Long, J., Pomerantz, R.,
 Ioshikhes, I.P., Shim, C., Decker, J., Thomas, E., Ferreira, A.,
 Gordon, M., Goltz, J.S. and Kucherlapati, R.
 High Throughput Mouse Sequencing
 Unpublished

TITLE
 JOURNAL 2 (bases 1 to 254993)
 REFERENCES
 Montgomery, K.T., Grills, G., Han, J., Lee, E., Long, J., Pomerantz, R.,
 Ioshikhes, I.P., Shim, C., Decker, J., Thomas, E., Ferreira, A.,
 Gordon, M., Goltz, J.S. and Kucherlapati, R.
 Direct Submission
 Submitted (08-AUG-2000) Department of Molecular Genetics, Albert
 Einstein College of Medicine Genome Center, 1300 Morris Park Ave.,
 Bronx, NY 10461, USA
 On Jan 26, 2002 this sequence version replaced gi:14488282.

COMMENT

-----Genome Center
 Center: Harvard Partners Genome Center
 Center Code: HPGC
 Web site: <http://www.hpcg.org/Sequence/mouse.html>
 Contact: hpcg@mednet.mgh.harvard.edu
 -----Summary Statistics
 Center project name: MAP
 Sequencing vector: pUC18, L08752
 Chemistry: Dye-terminator Big Dye, 100%
 *Consensus quality: 247096 at least Q20
 *Consensus quality: 242518 at least Q30

*Consensus quality: 235589 at least Q40
*Estimated insert size: agarose-FP - N/A
*Estimated insert size: 254133 - sum-of-contigs
Quality coverage: agarose-FP - N/A
Quality coverage: 6 x in Q20 bases; sum-of-contigs estimation

* NOTE: This is a 'working draft' sequence. It currently
* consists of 44 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 27415: contig of 27415 bp in length
* 27416 27435: gap of unknown length
* 27436 57817: contig of 30382 bp in length
* 57818 57837: gap of unknown length
* 57838 81637: contig of 23800 bp in length
* 81638 81657: gap of unknown length
* 81658 102155: contig of 20498 bp in length
* 102156 102175: gap of unknown length
* 102176 121200: contig of 19025 bp in length
* 121201 121220: gap of unknown length
* 121221 135073: contig of 13853 bp in length
* 135074 135093: gap of unknown length
* 135094 156689: contig of 21596 bp in length
* 156690 156709: gap of unknown length
* 156710 166185: contig of 9476 bp in length
* 166186 166205: gap of unknown length
* 166206 178798: contig of 12593 bp in length
* 178819 187247: contig of 8429 bp in length
* 187248 187267: gap of unknown length
* 187268 192492: contig of 5225 bp in length
* 192493 192512: gap of unknown length
* 192513 200084: contig of 7572 bp in length
* 200085 200104: gap of unknown length
* 200105 207054: contig of 6950 bp in length
* 207055 207074: gap of unknown length
* 207075 212944: contig of 5870 bp in length
* 212945 219664: gap of unknown length
* 219665 219468: contig of 6504 bp in length
* 219469 219488: gap of unknown length
* 219489 220369: contig of 881 bp in length
* 220370 220389: gap of unknown length
* 220390 221164: contig of 775 bp in length
* 221165 221184: gap of unknown length
* 221185 222796: contig of 1612 bp in length
* 222797 222816: gap of unknown length
* 222817 224762: contig of 1946 bp in length
* 224763 224782: gap of unknown length
* 224783 225999: contig of 1217 bp in length
* 226000 226019: gap of unknown length
* 226020 227409: contig of 1390 bp in length
* 227410 227429: gap of unknown length
* 227430 230145: contig of 2716 bp in length
* 230146 230165: gap of unknown length
* 230166 231842: contig of 1677 bp in length
* 231843 231862: gap of unknown length
* 231863 232786: contig of 924 bp in length
* 232787 232806: gap of unknown length
* 232807 233137: contig of 331 bp in length
* 233138 233157: gap of unknown length
* 233158 234519: contig of 1362 bp in length
* 234520 234539: gap of unknown length
* 234540 236416: contig of 1877 bp in length
* 236417 236436: gap of unknown length
* 236437 237967: contig of 1531 bp in length
* 237968 237987: gap of unknown length
* 237988 239220: contig of 1233 bp in length
* 239221 239240: gap of unknown length
* 239241 239367: contig of 127 bp in length

* 239368 239387: gap of unknown length
* 239388 239395: contig of 572 bp in length
* 239396 239979: gap of unknown length
* 239980 241016: contig of 1037 bp in length
* 241017 241036: gap of unknown length
* 241037 242391: contig of 1355 bp in length
* 242392 242411: gap of unknown length
* 242412 243823: contig of 1412 bp in length
* 243824 243843: gap of unknown length
* 243844 245227: contig of 1384 bp in length
* 245228 245247: gap of unknown length
* 245248 246334: contig of 1087 bp in length
* 246335 246354: gap of unknown length
* 246355 247451: contig of 1097 bp in length
* 247452 247471: gap of unknown length
* 247472 248780: contig of 1309 bp in length
* 248781 248800: gap of unknown length
* 248801 250041: contig of 1241 bp in length
* 250042 250061: gap of unknown length
* 250062 251173: contig of 1112 bp in length
* 251174 251193: gap of unknown length
* 251194 252713: contig of 1520 bp in length
* 252714 252733: gap of unknown length
* 252734 254014: contig of 1281 bp in length
* 254015 254034: gap of unknown length
* 254035 254176: contig of 142 bp in length
* 254177 254196: gap of unknown length
* 254197 254993: contig of 797 bp in length.
Location/Qualifiers
1. .254993
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL6/J"
/db_xref="taxon:10090"
/chromosome="6"
/clone="RP23-117I23"
/sex="male"
1. .27415
/note="assembly_name:Contig138"
27436. .57817
/note="assembly_name:Contig137
clone_end:SP6
vector_side:left"
57838. .81637
/note="assembly_name:Contig136"
81658. .102155
/note="assembly_name:Contig135"
102176. .121200
/note="assembly_name:Contig134"
121221. .135073
/note="assembly_name:Contig133"
135094. .156689
/note="assembly_name:Contig132"
156710. .166185
/note="assembly_name:Contig131
clone_end:T7
vector_side:left"
166206. .178798
/note="assembly_name:Contig130"
178819. .187247
/note="assembly_name:Contig129"
187268. .192492
/note="assembly_name:Contig128"
192513. .200084
/note="assembly_name:Contig127"
200105. .207054
/note="assembly_name:Contig126"
207075. .212944
/note="assembly_name:Contig125"
212965. .219468
/note="assembly_name:Contig124"
219489. .220369
/note="assembly_name:Contig123"

```

misc_feature      220390..221164
                  /note="assembly_name:Contig122"
misc_feature      221185..222796
                  /note="assembly_name:Contig121"
misc_feature      222817..224762
                  /note="assembly_name:Contig120"
misc_feature      224783..225999
                  /note="assembly_name:Contig119"
misc_feature      226020..227409
                  /note="assembly_name:Contig118"
misc_feature      227430..230145
                  /note="assembly_name:Contig117"
misc_feature      230166..231842
                  /note="assembly_name:Contig116"
misc_feature      231863..232786
                  /note="assembly_name:Contig115"
misc_feature      232807..233137
                  /note="assembly_name:Contig114"

```

Alignment Scores:

```

Pred. No.:      1.44e-37      Length:      254993
Score:          45.00         Matches:      77
Percent Similarity: 97.47%     Conservative: 0
Best Local Similarity: 97.47%   Mismatches:    1
Query Match:     16.25%        Indels:        2
DB:              2           Gaps:          0

```

US-09-896-522-2 (1-277) x AC078885 (1-254993)

```

QY      67  PheTYrLySVaIleuThrAlaGluGlnLyAlaLySAlaLeuLySGInTYrAsnPh 86
          |||||
Db      156571 TTCTACAGGTTCTGACGGCGACAGCAGGCTTTGAAGGACAGTACAAATTT 156512
          |||||

QY      87  AspHisProAspAlaPheAspAsnAspLeuMetHisArg-ThrLeuLysAsnIleValG 106
          |||||
Db      156511 GACCACCACAGATGCTTTGATTAATGATGATGATGACACAA-GACCTGAAAAAACATTGTTGA 156453
          |||||

QY      106 uGIyLySVThValGluValProThrTYrAspPheValThrHisSerArgLeuProGIuTh 126
          |||||
Db      156452 AGGCAAAACTGTGAGGTCCTACTATGATTTTGTGACCCACTCAAGGTTTACCAGAGAC 156393
          |||||

QY      126 rThrValValTYrProAlaAspValValLeuPheGluGlyIleLeuValPheTYr 144
          |||||
Db      156392 CACTGTGTGTACCACTGATGTGTCTGTTCGAGGCACTTGTGATTCTAC 156338
          |||||

```

Search completed: November 25, 2003, 11:28:57
 Job time : 4109 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 25, 2003, 08:11:01 ; Search time 2545 Seconds

(without alignments)
2645.320 Million cell updates/sec

Title: US-09-896-522-2
Perfect score: 1450
Sequence: 1 MASAGGEDCSPPAEADRP.....HPGMLTSGKSHLESSSRPH 277

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 22781392 segs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seg length: 0
Maximum DB seg length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=x1h
-Q=/cgr2_1/USPRO.spool/US09896522/runat_21112003_184354_5011/app_query.faeta_1.455
-DB=EST -QMT=fastlap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45
-DOCLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pro -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USFR=US09896522.@CGN_1.1_2810.@runat_21112003_184354_5011 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6
-Fgapext=7 -Ygapop=10 -Ygapext=0.5 -Delop=6 -Delext=7

Database :

EST: *
1: em_estba: *
2: em_estbm: *
3: em_estin: *
4: em_estnu: *
5: em_estov: *
6: em_estpl: *
7: em_estro: *
8: em_hic: *
9: gb_est1: *
10: gb_est2: *
11: gb_hic: *
12: gb_est3: *
13: gb_est4: *
14: gb_est5: *
15: em_estfun: *
16: em_estom: *
17: em_gss_hum: *
18: em_gss_inv: *
19: em_gss_pln: *
20: em_gss_vit: *
21: em_gss_fun: *
22: em_gss_mam: *
23: em_gss_mus: *
24: em_gss_pro: *
25: em_gss_rtd: *
26: em_gss_phg: *
27: em_gss_vrl: *
28: gb_gss1: *

29: gb_gss2: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1401	96.6	1118	12	BM467984 AGENCOURT
2	1346	92.8	1201	13	BX394295 BX394295
3	1332	91.9	1114	13	BX343101 BX343101
4	1327	91.5	998	13	BX400889 BX400889
5	1316	90.8	916	12	BG826894 BG2750978
6	1314	90.6	860	10	BG491358 BG491358
7	1313	90.6	1051	12	BM459108 BM459108
8	1311.5	90.4	869	10	BG491384 AGENCOURT
9	1300.5	89.7	848	12	B1258532 B1258532
10	1278	88.1	1044	12	BM917506 BM917506
11	1277	88.1	956	13	BQ719741 BQ719741
12	1274	87.9	943	13	BQ511803 BQ511803
13	1270	87.6	1036	13	BQ072501 BQ072501
14	1260	86.9	961	12	B1416102 B1416102
15	1259	86.8	922	10	BG390519 BG390519
16	1257	86.7	922	13	BQ935919 BQ935919
17	1246	85.9	827	12	B1261258 B1261258
18	1238	85.4	810	9	B1753007 B1753007
19	1232	85.0	779	12	AU131406 AU131406
20	1226	84.6	819	12	B1910015 B1910015
21	1224	84.4	725	12	B1256928 B1256928
22	1219.5	84.1	959	12	B1770572 B1770572
23	1215	83.7	796	9	AU141725 AU141725
24	1199	82.8	889	13	BQ885655 BQ885655
25	1195	82.4	898	14	CD514811 CD514811
26	1188.5	82.0	819	12	B1770639 B1770639
27	1183	81.6	822	12	BM948090 BM948090
28	1168.5	80.6	910	13	BQ437115 BQ437115
29	1132	78.1	952	10	BG163897 BG163897
30	1119	77.2	812	12	BG974826 BG974826
31	1115.5	76.9	1055	12	BM545603 BM545603
32	1113.5	76.8	875	13	BQ714585 BQ714585
33	1107	76.3	772	10	BF346907 BF346907
34	1105	76.2	730	12	BG770518 BG770518
35	1096	75.6	768	13	BQ614140 BQ614140
36	1083.5	74.7	876	13	BQ292179 BQ292179
37	1074.5	74.1	753	9	AU122008 AU122008
38	1069	73.7	778	14	CD101693 CD101693
39	1068.5	73.7	734	14	CA386173 CA386173
40	1058	72.7	1039	12	BM803129 BM803129
41	1054	72.7	1095	12	BM802939 BM802939
42	1041.5	71.8	912	12	B1412715 B1412715
43	1030	71.0	686	12	B1526272 B1526272
44	1030	71.0	694	12	B1906603 B1906603
45	1009.5	69.6	863	12	B1085319 B1085319

ALIGNMENTS

RESULT 1
LOCUS BM467984 1118 bp mRNA linear EST 05-FEB-2002
DEFINITION AGENCOURT_6437937 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:5532885
5', mRNA sequence.
ACCESSION BM467984
VERSION BM467984.1 GI:18517026
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo.
REFERENCE 1 (bases 1 to 1118)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC

CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
Plate: L1AM12216 row: n column: 22
High quality sequence stop: 689.
Location/Qualifiers

FEATURES
source 1..1118
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5532885"
/issue_type="leiomysarcoma"
/lab_host="DH10B (phage-resistant)"
/clone_1ib="NIH MGC 71"
/note="Organ: uterus; Vector: PCMV-SPORE; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 2.1 kb."
BASE COUNT 234 a 311 c 361 g 209 t 3 others
ORIGIN

Alignment Scores:

Pred. No.: 4.72e-157 Length: 1118
Score: 1401.00 Matches: 271
Percent Similarity: 98.1% Conservative: 1
Best Local Similarity: 97.83% Mismatches: 4
Query Match: 96.62% Indels: 1
Gaps: 0

US-09-896-522-2 (1-277) x BM467984 (1-1118)

QY 1 MetAlaSerAlaGlyGlyGluAspCysGluSerProAlaProGluAlaAspArgProHis 20
Db 47 ATGGCTTCGGCGGAGCGGAGAGACTCGAGAGCCCGCGCGGAGCGGAGCGGAGCGGAGC 106
QY 21 GlnArgProPheLeuIleGlyValSerGlyGlyThrAlaSerGlyValSerThrValCys 40
Db 107 CAGCGGCCCTTCGATAGGGGTAGCGCGGCGGAGCTGCACGGGAACTCCAGCTGTGT 166
QY 41 GlnValIleMetGluLeuLeuGlyGlnAsnGluValGluGlnArgGlnArgValVal 60
Db 167 GAGAAAGATCATGAGATTCCTGGGACAGACAGAGTGGAAACGCGGACCGGAGAGGTGTC 226
QY 61 IleLeuSerGlnAspArgPheTyrllyValIleuThralaGluGlnValAlaValAlaLeu 80
Db 227 ATCTGAGCCAGGACAGGCTTCTACAGAGTCTTGAAGCGGAGAGGAGGAGGAGGCTTG 286
QY 81 LysGlyGlnIlyrAsnPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThr 100
Db 287 AAAGACAGTCAATTGACCATTCAGATCCCTTGTATTAATGATTGATGACACAGGACT 346
QY 101 LeuValAsnIleValGluGlyLysThrValGluValProThrTyrrAspPheValThrHis 120
Db 347 CTGAAGAAATCATCTGGAGGCGGAAACGAGTGGAGCGGAGCTATGATTTGTGACACAC 406
QY 121 SerArgLeuProGluThrThrValValTyrrProAlaAspValValLeuPheGluGlyIle 140
Db 407 TCAAGGTACCAAGACACAGCGGTGTACCTCGGAGCGGTCTTGTGAGGAGCATC 466
QY 141 LeuValPheTyrrSerGlnGluIleArgAspMetPheHisLeuAlaGluPheValAspThr 160
Db 467 TTGGTGTCTACAGCAGGAGATCCGGGACATGTTCCACCTGCGCTTGTGTGACACCC 526
QY 161 AspSerAspValArgLeuSerArgArgValLeuArgAspValArgArgGlyValArgAspLeu 180

Db 527 GACTCCGAGCTACGGCTGTCTCGAAGATTCTCCGGAGAGTGGCCGAGAGGAGGACTG 586
QY 181 GlnGlnIleLeuThrGlnIlyrThrPheValLysProAlaPheGluGluPheCysLeu 200
Db 587 GAGCAGATTCGACGAGGAGTACACACACCTTCGTGTAAACCGGCTTCCAGAGATTCCTCC 646
QY 201 ProThrLysLysThrAlaAspValIleIleProArgGlyValAspAsnMetValAlaIle 220
Db 647 CCACAAAGAGATGCGCATGTGATCATCCACGAGAGTGGACATATGTTGCCATC 706
QY 221 AsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysLysTrpHisArg 240
Db 707 AACCTATGCTGACGACACATCCAGGACATTTCTGATGATGATGATGATGATGATGATG 766
QY 241 GlyGlySerAsnGlyArgSerTyrllyValGlnPheSerGluProGlyAspHisProGly 260
Db 767 GGAGGGTCCAAATGGGCGGAGCTCAAGCGGACCTTTCTGAGGACGAGGAGCCCTGGN 826
QY 261 -MetLeuThrSerGlyLysArgSerHisLeuGluSerSerArgPro 276
Db 827 GATGCTGACCTCTGGCAACGCTCACAATTGTGAGTCCAGCAGCAGCC 875

RESULT 2
BX394295 1201 bp mRNA linear EST 13-MAY-2003
LOCUS BX394295 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
DEFINITION BX394295.1 GI:30624219
ACCESSION BX394295
VERSION BX394295.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1201)
AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segre@genoscope.cns.fr, Web: www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 4968.r For more information about this cluster, see <http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0DC013AC070P1&cluster=4968.r>. Contact: Peng Liang Email: filang@life.techn.com URL: <http://fulllength.invitrogen.com/InvitrogenCorporation> 1600 Faraday Avenue Genoscope sequence ID: CS0DC013AC070P1.

FEATURES

source 1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DC013YE13"
/issue_type="NEUROBLASTOMA COT 25-NORMALIZED"
/note="Left strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT 288 a 284 c 363 g 226 t 40 others
ORIGIN

Alignment Scores:

Pred. No.: 2.09e-150 Length: 1201
Score: 1346.00 Matches: 270
Percent Similarity: 98.55% Conservative: 2
Best Local Similarity: 97.83% Mismatches: 4
Query Match: 92.83% Indels: 3
Gaps: 0

US-09-896-522-2 (1-277) x BX34295 (1-1201)

QY 1 MetAlaSerAlaGlyGlyAlaAspCysGluSerProAlaProGluAlaAspArgProHis 20
 DB 88 ATGGCTTCGGCGGGAGCGGAGACTGCGAGAGCCCCCGCGGAGCC-GACCGTCCGAC 146
 QY 21 GlnArgProPheLeuIleGlyValSerGlyGlyThrAlaSerGlyLysSerThrValCys 40
 DB 147 CAGGGGCGCTTCCTGATAGGGGTGAGCGCGGACCTCCAGCGGGAGAGTGCACCGTGTGT 206
 QY 41 GlnLysIleMetGluLeuLeuGlyGlnAsnGluValGluGlnArgGlnArgValVal 60
 DB 207 GAGAAAGATCATGAGGTGCTGGGACGAAAGAGGTGAAACAGCGCGCAAGAGTGTCT 266
 QY 61 IleLeuSerGlnAspArgPheThrLysValLeuThrAlaGluGlnLysAlaLysAlaLeu 80
 DB 267 ATCTGAGCCAGGACAGGCTTCTACAGAGTCTTGACGAGGAGCAAGAGCCAGGCTTG 326
 QY 81 LysGlyGlnThrAsnPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThr 100
 DB 327 AAGGACAGATCAATTTTGAACATCCAGATGCTTTGATATGATTTGATGCACAGACT 386
 QY 101 LeuLysAsnIleValGluGlyLysThrValGluValProThrTyraAspPheValThrHis 120
 DB 387 CTAAAGAACATCTGTGAGGGGCAAAACGCTGAGGTGCGCACTATGATTTTGTGACACAC 446
 QY 121 SerArgLeuProGluThrThrValValTyraProAlaAspValValLeuPheGluGlyIle 140
 DB 447 TCAGGTTACAGAGACGACGAGGTGTCTACCTGCGGACGTTGCTGTTGAGGGGATC 506
 QY 141 LeuValPheTyraSerGlnGluIleArgAspMetPheHisLeuArgLeuPheValAspThr 160
 DB 507 TTGGTGTTCACAGCCAGGAGATCCGGGACATGTTCCACCTGCGCTCTTCTGTGACACC 566
 QY 161 AspSerAspValArgLeuSerArgValLeuArgAspValArgArgGlyArgAspLeu 180
 DB 567 GACTCCGACCTCAGGCTGTCTCGAAGAGTCTCCGGGACCTGGCGGAGGAGGAGACTTG 626
 QY 181 GluGlnIleLeuThrGlnTyraThrThrPheValLysProAlaPheGluGluPheCysLeu 200
 DB 627 GAGCAGATTCAGGACGATCACACCTTGTAAGCGGCTTTCGAGGAGTCTGCTG 686
 QY 201 ProThrLysLysTyraAlaAspValIleIleProArgGlyValAlaAspMetValAlaIle 220
 DB 687 CCAGCAAGAAGATGCGCATGATCATCCACGAGAGTGCACATATGCTTGCATC 746
 QY 221 AsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysLysThrHisArg 240
 DB 747 AACCTGATCGTGACGACATCCAGGACATTCGAAATGTGACATCTTCAAAATGGACCGA 806
 QY 241 GlnGlySerAsnGlyArgSerTyraArgThrPheSerGluProGlyAspHisProGly 260
 DB 807 GGGGGGTCCATGGCGGAC-TACAGCGGACCTTTTCTAGCGGAGGGA-CACCTGGG 864
 QY 261 MetLeuThrSerGlyLysArgSerHisLeuGluSerSerArgPro 276
 DB 865 ATCTGATCTCTGCGCAACGCTACATTTGAGTCAAGACAGACCA 912

RESULT 3
 LOCUS BX343101 1114 bp mRNA linear EST 02-MAY-2003
 DEFINITION BX343101 Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED
 Homo sapiens cDNA clone CS0DL009YD15 5-PRIME, mRNA sequence.
 ACCESSION BX343101
 VERSION BX343101.1 GI:30334170
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 1114)
 AUTHORS Li, W.-B., Gruber, C., Jessee, J. and Polyes, D.
 TITLE Full-length cDNA libraries and normalization

JOURNAL COMMENT Unpublished
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 Evry cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 4968.r For
 more information about this cluster, see
 http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CS0DL009CB080P1&cluster=4968.r. Contact :
 Feng Liang Email : fliang@lifetech.com URL :
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
 Faraday Avenue Genoscope sequence ID : CS0DL009CB080P1.

FEATURES
 source
 1..1114
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DL009YD15"
 /cell_type="B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"
 /cell_line="RAMOS CELL LINE"
 /clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

BASE COUNT 247 a 291 c 360 g 202 t 14 others
 ORIGIN

Alignment Scores:
 Pred. No.: 8,89e-149 Length: 1114
 Score: 1332.00 Matches: 268
 Percent Similarity: 97.83% Conservative: 2
 Best Local Similarity: 97.10% Mismatches: 6
 Query Match: 91.86% Indels: 3
 DB: 13 Gaps: 0

US-09-896-522-2 (1-277) x BX343101 (1-1114)

QY 1 MetAlaSerAlaGlyGlyAlaAspCysGluSerProAlaProGluAlaAspArgProHis 20
 DB 159 ATGGCTTCGGCGGGAGCGGAGACTGCGAGAGCCCCCGCGGAGSC-GACCGTCCGAC 217
 QY 21 GlnArgProPheLeuIleGlyValSerGlyGlyThrAlaSerGlyLysSerThrValCys 40
 DB 218 CAGCGGCCCTTCCTGATAGGGGTGAGCGCGGACCTCCAGCGGGAGAGTGCACCGTGTGT 277
 QY 41 GlnLysIleMetGluLeuLeuGlyGlnAsnGluValGluGlnArgGlnArgValVal 60
 DB 278 GAGAGATCATGAGTGTCTGGGACGAAAGAGTGAACAGCGGCGGAGGAGGGTGTCT 337
 QY 61 IleLeuSerGlnAspArgPheThrLysValLeuThrAlaGluGlnLysAlaLysAlaLeu 80
 DB 338 ATCTGAGCCAGGACAGGCTTCTACAGAGTCTTGACGAGGACAGACAGGCGCAAGGCTTG 397
 QY 81 LysGlyGlnThrAsnPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThr 100
 DB 398 AAGGACAGATCAATTTTGAACATCCAGATGCTTTGATATGATTTGATGCACAGACT 457
 QY 101 LeuLysAsnIleValGluGlyLysThrValGluValProThrTyraAspPheValThrHis 120
 DB 458 CTGAGAGATCATGAGGAGGCGCAAAACGCTGAGGTGCGCACTATGATTTTGTGACACAC 517
 QY 121 SerArgLeuProGluThrThrValValTyraProAlaAspValValLeuPheGluGlyIle 140
 DB 518 TCAGGTTACAGAGACCGAGTGTCTACCTGCGGAGTGTCTGTTGAGGGGATC 577
 QY 141 LeuValPheTyraSerGlnGluIleArgAspMetPheHisLeuArgLeuPheValAspThr 160
 DB 578 TTGGTGTTCACAGCGGAGGATTCGGGACAGTTCACCTGCGCTCTTCTGTGACACC 637
 QY 161 AspSerAspValArgLeuSerArgValLeuArgAspValArgArgGlyArgAspLeu 180

Db	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
Db	638	GACTCCGACGCTCAGGCTGTCTTCGAGAGATTCTCCGGAGCGTGCCTCCGAGGGAGGACCTG	697									
Qy	181	GluguinileuthrGlnTyrThrThrpheValIysProalaIpheGluGluIpheCysIleu	200									
Db	698	GAGCAGATTTCGACGCATTAACACCACTTCGTAAGCCGGCCTTCGAGAGATTTCGCTG	757									
Qy	201	ProthrIysIysIleTyrAlaAspValIleIleIleProIargIValAlaAspMetValAlaIle	220									
Db	758	CCGACAAAGAAAGTAAAGCCGATGTATCATCCACAGAGAGTGGACAAATATGTTGGCCATC	817									
Qy	221	AsnIleuIleValGlnHisIleIleGlnAspIleIleuAsnIysAspIleCysIysTrpHisArg	240									
Db	818	AACTGATTCGTGCACACATCCAGACATTCGTAATGTGTACATTCGAAATGGCACCGA	877									
Qy	241	GlyIysSerAsnIysArgSerTyrIysArgThrPheSerGluProGlyAspHisProGly	260									
Db	878	GGAGGTCCTCAATGGCGCGAG-C-TACAAAGGAGACCTTTCTTCGAGCCAGGGGA-CACCTGG	935									
Qy	261	MetIeuThrSerGlyIysArgSerHisIleuGluSerSerArgPro	276									
Db	936	ATGTGACCTCTGGCAAAAGCTCATTTGGAGTGCAGCAGACGACCA	983									
RESULT 4												
EX400889												
LOCUS	EX400889	Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens	998 bp	mRNA	linear	EST 13-MAY-2003						
DEFINITION	EX400889	CDNA clone CSODK003YL24 5-PRIME, mRNA sequence.										
ACCESSION	EX400889											
VERSION	EX400889.1	GI:30622359										
KEYWORDS	EST.											
SOURCE	Homo sapiens (human)											
ORGANISM	Homo sapiens											
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;											
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.											
	1 (bases 1 to 998)											
	Li, W.B., Gruber, C., Jessee, J. and Polayes, D.											
	Full-length cDNA libraries and normalization											
	Unpublished											
	Contact: Genoscope											
	Genoscope - Centre National de Sequencage											
	BP 191 91006 Evry cedex - France											
	Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr											
	Library was constructed by Life Technologies, a division of											
	Invitrogen. This sequence belongs to sequence cluster 4968.r for											
	more information about this cluster, see											
	http://www.genoscope.cns.fr/											
	cgf-bin/cluster.cgi?seq=CSODK003DP12P1&cluster=4968.r. Contact :											
	Peng Liang Email : lliang@lifetech.com URL :											
	http://fulllength.invitrogen.com/ Invitrogen Corporation 1600											
	Paraday Avenue Genoscope sequence ID : CSODK003DF120P1.											
	Location/Qualifiers											
FEATURES	source											
	1..998											
	/organism="Homo sapiens"											
	/mol_type="mRNA"											

DB:	13	Gaps:	0
US-09-896-522-2 (1-277) x BX400889 (1-998)			
OY	1	Metalase1agiylgiyuaspCygluserPro1a1aaspargProhis	20
Db	67	ATGGCTTCGGGGGAGGCGAAGACTGGAGAGCCCGCGCGGA-SCGNACCGTCCGAC	125
OY	21	GLnarGProPheleuileGlyValSerGlyThrAlaSerGlyLysSerThValCys	40
Db	126	CAGGAGCCCTTCCGATAGGGGTGAGGGCGGCGACTCCAGCGGGAATGCACGTGTGT	185
OY	41	GLuylsIleMetGluLeuLeuGlyGLAsnGluValGluGlnAspGlnAspLysValVal	60
Db	186	GAGAAAGATCATGAGCTTGCTGGGACAGAACAGGTGGAAACAGCGGACGGAGGTGGTC	245
OY	61	IleLeuSerGlnaSPargPheTyLysValIleuThra1agiunGlySalAlaLeu	80
Db	246	ATCTGTAGCCAGGACAGGTTCTAACAAGTCTGTAGCGGACAGGAGGCGCAAGGCTTG	305
OY	81	LysGlyGlnTyraenPheaspHisProaspAlaPheaspnaPLeuMetHisArgThr	100
Db	306	AAAGGACAGTGCATTTTGGACCAATCCAGATGCGCTTTGATATGATTGATGTCACAGACT	365
OY	101	LeuLysaenIleValGluGlyLysThValGluValProThrTyraSPheValThHis	120
Db	366	CTGAGAAACATCGTGGAGGGCGAAAACGATGAGGTGCGACATATGATTGTGACACAC	425
OY	121	SerArgLeuProGluThrThValValTyProAlaSPValValLeuPheGluGlyLe	140
Db	426	TCAAGGTTACAGAGACCAAGGTGTCTACCTCGGACGTGGTCTGTTGAGGGCATC	485
OY	141	LeuValPheTySerGlnGluIleArgaspMetPheHisLeuArgLeuPheValAspThr	160
Db	486	TTGGTGTCTTCACAGCCAGAGATCCGGGACATGTTTCAACCTGCGCTCTTGTCGACACC	545
OY	161	AspSeraspValaArgLeuSerArgArgValIleuArgaspValaArgArgGlyArgAspLeu	180
Db	546	GACCCGACGTCAGAGCGTGTCTCGAAGAGTCTCCGGACGTGCGCCGACGAGGAGGACCTG	605
OY	181	GluGlnIleLeuThrGlnTyThrThPheValLysProAlaPheGluGluPheCysLeu	200
Db	606	GAGCAGATTCGAGGCA-TACACCACTCTGCTGAAGCCGGCTTCGAGAGTTTGCTGCTG	664
OY	201	ProThrLysLysTyraIaaspValIleIleProArgGlyValaAspMetValAlaIle	220
Db	665	CCGACAAAGAAATGATGCGATGATCATCCACGAGAGTGACAAATATGTTGCCATC	724
OY	221	AsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysLysThrHisArg	240
Db	725	AACTGTATCGGACGACATCCAGAAATTCGTAATGATGCACTCCAAATGGCACCGGA	784
OY	241	GlyGlySerAsnGlyArgSerTyTyLysArgThPheSerGluProGlyAspHisProGly	260
Db	785	GAGGGGTCATAGGCGGAC-TACAAAGCGACCTTTTCTGAGCCAGGGGA-CACCTCGGG	842
OY	261	MetLeuThrSerGlyLysArgSerHisLeuGluSerSerSerArgProHis	277
Db	843	ATGCTGACCTTGCGCAACGCTGCATTTGGAGTCCAGCAGCAACCCAC	893
RESULT 5			
LOCUS	BG826894	916 bp	mRNA linear EST 22-MAY-2002
DEFINITION	602750978P1 NIH_MGC_17 Homo sapiens CDNA clone IMNGE:4903819 5',		
ACCESSION	BG826894		
VERSION	BG826894.1	GI:14174481	
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo. 1 (bases 1 to 916)		

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-r@mail.nih.gov
 Tissue Procurement: ATCC

CDNA Library Preparation: Ling Hong/Rubin Laboratory
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Plate: LLCM1802 row: k column: 20
 High quality sequence stop: 843.
 Location/Qualifiers

FEATURES

source
 1. 916
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4903819"
 /tissue_type="thadomyosarcoma"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC 17"
 /note="Organ: muscle; Vector: pOTB7; Site 1: EcoRI; Site 2: XhoI; CDNA made by oligo-dt priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size selected for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."
 BASE COUNT 203 a 239 c 309 g 165 t
 ORIGIN

Alignment Scores:

Pred. No.: 5,47e-147 Length: 916
 Score: 1316.00 Matches: 270
 Percent Similarity: 96.47% Conservative: 3
 Best Local Similarity: 95.41% Mismatches: 7
 Query Match: 90.76% Indels: 3
 DB: 12 Gaps: 0

US-09-896-522-2 (1-277) x BG826894 (1-916)

QY 1 MetAlSeraAaGlyGlyGluAaPcyGluSerProAlaProGluAlaAaPaPProHis 20
 DB 80 ATGGCTTCGGCGGAGCGGAGAGCTGCGAGAGCCCGCGCGGAGGCGCGCTCGGCAC 139
 QY 21 GlnAArgProPheLeuIleGlyValSerGlyGlyThraLaseGlyLysSerThrValCys 40
 DB 140 CAGCGGCGCTTCCTGATAGGAGGAGGAGGCGGCACTGCCAGCGGAGAGTGCAGCGTGT 139
 QY 41 GlnLysIleMetGluLeuLeuGlyGlnAaGlnValGluGlnAArgGlnArgValVal 60
 DB 200 GAGAAAGATCATGAGTTGCTGGAGCAAGAGAGTGGAAACAGCGGCGGAGAGTGTGTC 259
 QY 61 IleLeuSerGlnAaPArgPheTyLysValLeuThraLaseGluLysValValValLeu 80
 DB 260 ATCTTGAAGCAGACAGGCTCTCAAGGTCTTCAAGCGCAAGAGCAAGAGCGCTTG 319
 QY 81 LysGlyGlyLysPheAaPHisProAaPAlaPheAaPAsnAaPLeuMetHisArgThr 100
 DB 320 AAAAGGACAGTACATTTTGAACATCCAGATGCTTGTATATATATTTGATGACAGGACT 379
 QY 101 LeuLysAsnIleValGluGlyLysThrValGluValProThrTyraPheValThrHis 120
 DB 380 CTGAAGAACATCTGAGGCGCAAAACGCTGAGAGTGCAGCTATGATTTTGTGACACAC 439
 QY 121 SerArgLeuProGluThrThrValValTyProAlaAaPValValLeuPheGluGlyTle 140
 DB 440 TCAAGGTTCACAAAGACCAAGGAGTGTCTACCTTGCAGCACTGTGTTCTTGTGAGGCATC 499
 QY 141 LeuValPheTyraSerGlnGluIleArgAaPMetPheHisLeuArgLeuPheValAspThr 160

DB 500 TTGGTGTTCACAGCAGAGATCCGGAGCATGTTCCACTCGCTTCGTGGACACC 559
 QY 161 AaPSerAaPValArgLeuSerArgValLeuAaPValArgArgValArgValPleu 180
 DB 560 GACTCCGACGACAGGCTGTCTCGAAGAGTTCCTCCGGAGAGTCCGCGAGGAGGACCTG 619
 QY 181 GlnGlnIleLeuThraGlyThrThrPheValLysProAlaPheGluGluPheCysLeu 200
 DB 620 GAGCAGATTCACGACAGACACACTTCGTGAAGCCGCGCTTCAGAGAGTTCCTCG 679
 QY 201 ProThrLysTyraLysValAaPValIleLeuProArgValAaPAsnMetValAla1 220
 DB 680 CCGACAAAGATATGCGCATGTGATCATCCACAGAGAGTGGACAAATTTGTTGCCAT 739
 QY 220 eAaLeuIleValGlnHisIle-GlnAaPle-LeuAaGlyAaPLeuIleCysLysThrHis 239
 DB 740 CAACCTGATCGTGCAGCATTCACAGCATTCCTGAATGTGTGACATCTGCAAAATGGCAC 799
 QY 240 ArgGlyGlySerAaP-GlyAaPArgSerTyraLys-ArgThr-PheSerGluProGluAaPHis 258
 DB 800 CAGAGAGGATCCATAGGCGGAGGCTACAGAGCGGACCTTTTCTGAGACAGGAGAACAC 859
 QY 259 Pro-GlyMetLeuThrSerGlyLysArgSerHisLeuGluSerSerSerArgPro 276
 DB 860 CTTGGGAATGCTGACCTTCTGCAAAAGGTCCCATTTGAGTGTCAAGACAGACACC 914

RESULT 6
 BG491358 860 bp mRNA linear EST 27-MAR-2001
 BG491358 60235642F1 NIH_MGC_41 Homo sapiens CDNA clone IMAGE:4684738 5',
 DEFINITION mRNA sequence.
 ACCESSION BG491358.1 GI:13452870
 VERSION BG491358
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 1 (bases 1 to 860)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-r@mail.nih.gov
 Tissue Procurement: DCTD/DTF

CDNA Library Preparation: Ling Hong/Rubin Laboratory
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Plate: LLCM1493 row: c column: 11
 High quality sequence stop: 844.
 Location/Qualifiers

FEATURES

source
 1. 860
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4684738"
 /tissue_type="amelanotic melanoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC 41"
 /note="Organ: skin; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; CDNA made by oligo-dt priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH MGC library."
 BASE COUNT 195 a 223 c 288 g 154 t
 ORIGIN

Alignment Scores:

Pred. No.: 8.63e-147 Length: 860
 Score: 1314.00 Matches: 262
 Percent Similarity: 98.14% Conservative: 2
 Best Local Similarity: 97.40% Mismatches: 2
 Query Match: 90.62% Indels: 4
 DB: 10 Gaps: 0

US-09-896-522-2 (1-277) x BM459108 (1-860)

QY 1 MetAlaSerAlaGlyGlyGluAspCysGluSerProAlaProGluAlaAspArgProHis 20
 Db 52 ATGGCTTCGGCGGAGGCGGAAGATGCGAGAGCCCGCGCGGCGGAGCGGACCTGCCAC 111
 QY 21 GlnArgProPheLeuIleGlyValSerGlyGlyThrAlaSerGlyLysSerThrValCys 40
 Db 112 CAGCGGCCCTTCCTGATAGGGGTGAGCGCGCACTGCGACGGGAGAGTCCAGCTGTGT 171
 QY 41 GluLysIleMetGluLeuLeuGlyGlnAsnGluValGluGlnArgGlnArgLysVal 60
 Db 172 GAGAAAGATCATGAGTGTCTGGACAGAACGAGGTGGAACGCGGCGAGCGGAGAGTGGTC 231
 QY 61 IleLeuSerGlnAspArgPheTyrlLysValLeuThrAlaGluGlnLysAlaLysAlaLeu 80
 Db 232 ATCTTGAGCCAGGACAGGTTCTACAGAGTCTCGACGCGAGCAGAGGCGCAAGCCCTTG 291
 QY 81 LysGlyGlnIleTyrAsnPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThr 100
 Db 292 AAAGAGACATGATTTTGTACCATCCAGATCCCTTGTATTAATGATTTGATGCACAGACT 351
 QY 101 LeuLysAsnIleValGluGlyLysThrValGluValProThrTyrAspPheValThrHis 120
 Db 352 CTGAAGAACATCGTGAAGGCGAAACGGTGAAGTGGCGACTATGATTTTGTACACAC 411
 QY 121 SerArgLeuProGluThrThrValValTyrProAlaAspValValLeuPheGluGlyIle 140
 Db 412 TCAGAGTTACAGAGACACACGCTGTCTACCTCGGACGCTGTCTGTGAGGCGATC 471
 QY 141 LeuValPheTyrSerGlnGluIleArgAspMetPheHisLeuArgLeuPheValAspThr 160
 Db 472 TTGGTGTCTTACAGCCAGAGATCCGGGACATGTTCCACCTGCCCTCTCTGTGAGACAC 531
 QY 161 AspSerAspValArgLeuSerArgArg-ValLeuArgAspValArgArgGlyArgAspLe 180
 Db 532 GACTCCGACGTCAGGCTGTCTCGAAGAAGTCTCGGAGCGTGGCGGAGGAGGAGCCT 591
 QY 180 uGluGlnIleLeuThrGlnIleTyrThrPheValLysProAlaPheGluGluPheCysLe 200
 Db 592 GGAACATTCCTGACGAGTACACACCTTCGTGAAACCGGCTTCGAGGAGTCTGCTCCT 651
 QY 200 uProThrLysIleTyrAlaAspValIleIleProArgGlyValAspAsn-MetValAlaI 220
 Db 652 GCCGACAAAGAGATGCCCATGTGATCATCCGAGAGGTGACAAATATGTTGCCA 711
 QY 220 LeuLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysLysIleHis- 239
 Db 712 TCAACCTGATCGTGCAGACATCCAGACATTCGATGATGATGATCATCGAAATGCGACC 771
 QY 240 ArgGlyLysSerArgGlyArgSerTyrLysArgThrPheSerGluProGlyAspHisPro 259
 Db 772 CGAGAGAGTCCCATGAGGCGAGGTGACAGAGGAGCCTTTCT-GAGCGAGGAGGACCACT 830
 QY 260 GlyMetLeuThrSerGlyLys 266
 Db 831 GGGATGCTGACCTTGGGCAA 851

RESULT 7
 BM459108 1051 bp mRNA linear EST 05-FEB-2002
 LOCUS AGENCOURT_6414543 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5557433
 DEFINITION 5', mRNA sequence.
 ACCESSION BM459108

VERSION BM459108.1 GI:18508148
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 1051)
 AUTHORS NIH-MGC <http://mgs.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabs-remail.nih.gov
 Tissue Procurement: ATCC/DCPD/DRP
 CDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>
 Plate: LLAM12278 row: m column: 18
 Location/Qualifiers

FEATURES

source

1..1051
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5557433"
 /tissue_type="melanotic melanoma"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_72"
 /note="Organ: skin; Vector: pCMV-Sport6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 2 kb. Library constructed by Life Technologies."

BASE COUNT 237 a 264 c 312 g 221 t 17 others
 ORIGIN

Alignment Scores:

Pred. No.: 1.55e-146 Length: 1051
 Score: 1313.00 Matches: 258
 Percent Similarity: 94.57% Conservative: 3
 Best Local Similarity: 93.48% Mismatches: 13
 Query Match: 90.55% Indels: 2
 DB: 12 Gaps: 0

US-09-896-522-2 (1-277) x BM459108 (1-1051)

QY 1 MetAlaSerAlaGlyGlyGluAspCysGluSerProAlaProGluAlaAspArgProHis 20
 Db 6 ATGGCTTCGGCGGAGGCGGAAGATGCGAGAGCCCGCGGAGGCGGACCTGCCAC 65
 QY 21 GlnArgProPheLeuIleGlyValSerGlyGlyThrAlaSerGlyLysSerThrValCys 40
 Db 66 CAGCGGCCCTTCCTGATAGGGGTGAGCGCGGCACTGCGAGAGGAGAGTCCAGCTGTGT 125
 QY 41 GluLysIleMetGluLeuLeuGlyGlnAsnGluValGluGlnArgGlnArgLysVal 60
 Db 126 GAAAGATCATGAGTGTCTGGGACAGAACGAGGTGGAACAGGCGGAGGAGAGTGGTC 185
 QY 61 IleLeuSerGlnAspArgPheTyrlLysValLeuThrAlaGluGlnLysAlaLysAlaLeu 80
 Db 186 ATCTTGAGCCAGGACAGGTTCTTACAGAGTCTCGACGCGAGACAGAGGCGCAAGCTTG 245
 QY 81 LysGlyGlnIleTyrAsnPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThr 100
 Db 246 AAAGAGACATGATTTTGTACCATCCAGATGCCCTTGTATTAATGATTTGATGCACAGACT 305
 QY 101 LeuLysAsnIleValGluGlyLysThrValGluValProThrTyrAspPheValThrHis 120
 Db 306 CTGAAGAACATCGTGAAGGCGAAACGGTGAAGTGGCGACTATGATTTTGTGACACAC 365
 QY 121 SerArgLeuProGluThrThrValValTyrProAlaAspValValLeuPheGluGlyIle 140

Db 366 TCAAGTTACAGACACGAGTGTCTACCTTCGCGACGTTGTTCTGTTAGAGGCATC 425
 Qy 141 LeuValPheTyrSerGlnGluIleArgAspMetPheHisLeuArgLeuPheValAspThr 160
 Db 426 TTGGTGTCTACAGTCGAGATCCGGACATGTTCCACCTCGCCCTTCGTGAGACACC 485
 Qy 161 AspSerAspValArgLeuSerArgArgValLeuArgAspValArgArgGlyArgAspLeu 180
 Db 486 GACTCCGACATCAGAGATGCTCCAGAGATTCTCCGGACGCTGCTCCAGGAGGAGACTG 545
 Qy 181 GluGlnIleLeuThrGlnTyrThrThrPheValIleProAlaPheGluGluPheCysLeu 200
 Db 546 GACGACATTCGACGCGATCACACCTTCGTGAGTCGGCTTCCTCGAGAGTTCTGCTG 605
 Qy 201 ProThrIleValTyrAlaAspValIleIleProArgGlyValAspAsnMetValAlaIle 220
 Db 606 CCGACAAAGAGATGTCGATGTGATATCCACGAGAGTGGACATTAAGCTGACATC 665
 Qy 221 AsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysLysTrpHisArg 240
 Db 666 AACCTGATCCTGCAGTACATCCAGACATTTCTGAATGTGACATCTCAATGTGACCT 725
 Qy 241 GlyGlySerAsnGlyArgSerTyrLys-ArgThrPheSerGluProGlyAspHis-Arg 260
 Db 726 AGAGGTCATGCGGCGGAGCTACAGTCCGACCTTTTTCGAGTCAAGGAGCACCCCTG 785
 Qy 260 LysMetLeuThrSerGlyLysArgSerHisLeuGluIleSerSer 274
 Db 786 GGATGCTGACCTCGAAGAACGGTTCCTTTGGAGTTCACAT 829
 RESULT 8
 BG491384 869 bp mRNA linear EST 27-MAR-2001
 LOCUS BG491384 60235670Fl NIH_MGC_41 Homo sapiens cDNA clone IMAGE:4684936 5',
 DEFINITION mRNA sequence.
 ACCESSION BG491384
 VERSION BG491384.1 GI:13452896
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo.
 1 (bases 1 to 869)
 NIH-MGC http://mgs.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished
 Contact: Robert Strausberg, Ph.D.
 Email: csaps-r@mail.nih.gov
 Tissue Procurement: DCTD/BTP
 CDNA Library Preparation: Ling Hong/Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: LNC493 row: k column: 17
 High quality sequence stop: 863.
 Location/Qualifiers
 1..869
 source
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4684936"
 /tissue_type="amelanotic melanoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_id="NIH_MGC_41"
 /note="Organ: skin; Vector: pOTB7; Site 1: XhoI; Site 2:
 EcoRI; cDNA made by oligo-dT priming. Directionally cloned
 into EcoRI/XhoI sites using the following 5' adaptor:
 GGACGAG(G). Library constructed by Ling Hong in the
 laboratory of Gerald M. Rubin (University of California,
 Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
 Superscript II RT (Life Technologies). Note: this is a

BASE COUNT 197 a 223 c 291 g 156 t 2 others
 ORIGIN
 Alignment Scores:
 Pred. No.: 1,75e-146 Length: 869
 Score: 1311.50 Matches: 264
 Percent Similarity: 96.70% Conservat: 0
 Best local Similarity: 96.70% Mismatches: 4
 Query Match: 90.45% Indels: 5
 DB: 10 Gaps: 1
 US-09-896-522-2 (1-277) x BG491384 (1-869)
 Qy 1 MetAlaSerAlaGlyGlyGluAspCysGluSerProAlaProGluAlaAspArgProHis 20
 Db 52 ATGGCTTCGCGGAGCGAAGCTGCGAGAGCCCGCCCGGAGGCCACCTCCGCAC 111
 Qy 21 GlnArgProPheLeuIleGlyValSerGlyGlyThrAlaSerGlyLysSerThrValCys 40
 Db 112 CAGCGCCCTTCCTGTAAGGGGTGAGCGCGGACCTGCCAGCGGAAGTGCACCTGTGT 171
 Qy 41 GlyLysIleMetGluLeuLeuGlyGlnAsnGluValGluGlnArgGlnArgLysVal 60
 Db 172 GAGAAATCATGAGATTGCTGGGACAGAACGAGGTGAACAGCGCGCAGGAAGTGTGTC 231
 Qy 61 IleLeuSerGlnAspArgPheTyrLysValLeuThrAlaGluGlnLysAlaLysAlaLeu 80
 Db 232 ATCCTGAGCGAGACAGGTTCTACAGAGTCTCGACGCGAGAGAGCAAGGCGCACTTG 291
 Qy 81 LysGlyGlnTyrAsnPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThr 100
 Db 292 AAGGACAGTACATTTTGAACATCCAGATGCTTGATATGATTTGATGACAGAGACT 351
 Qy 101 LeuLysAsnIleValGluGlyLysThrValGluValProThrTyrAspPheValThrHis 120
 Db 352 CTGAAGAACATCGTGAAGGCAAAACGGTGAAGTCCGACCTATGATTTTGTGACACAC 411
 Qy 121 SerArgLeuProGluThrThrValValTyrProAlaAspValValLeuPheGluGlyIle 140
 Db 412 TCAAGTTACAGAGACCAAGGAGTCTACCCCTCGAGAGTGTGTTGTTGAGGCGATC 471
 Qy 141 LeuValPheTyrSerGlnGluIleArgAspMetPheHisLeuArgLeuPheValAspThr 160
 Db 472 TTGGTGTCTACAGCCAGAGATCCGGACATTTCCACTGCGCTTCCTCGAGACACC 531
 Qy 161 AspSerAspValArgLeuSerArgArgValLeuArgAspValArgArgGlyArgAspLeu 180
 Db 532 GACTCCGACGTACGGCTGTCTGGAAGATTCTCCGGAGCGTGGCGAGGAGGAGACTG 591
 Qy 181 GluGlnIleLeuThrGlnTyrThrPheValLysProAlaPheGluGluPheCysLeu 200
 Db 592 GAGCAATTCGACGAGATCACCACTTCGTMAACCGGCTTCAGAGGATTCCTGCTG 651
 Qy 201 ProThrLys-LysTyrAlaAspValIleIleProArgGlyValAspAsnMet-ValAlaI 220
 Db 652 CCGACAAAGAGATATGCGATGTGATCCACAGAGATGAGAACATATGAGTGTGCCA 711
 Qy 220 LeuAsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysLysTrpHis- 239
 Db 712 TCAACTGATCGTGGACGACATCAGACATTTCTGAATGTGACATCTGCAATGCGACACC 771
 Qy 240 ArgGlyGlySerAsn-GlyArgSerTyrLysArgThrPheSerGluProGlyAspHis-Arg 259
 Db 772 CGAGGAGGTCATGCGGCGGAGCTACAGCGGACCTTTCGAGCAGCAGACCA--CC 828
 Qy 259 ogLysMetLeuThrSerGlyLysArgSerHis 269
 Db 829 TGGAGTCTGACCTTCGCAACGATCAAT 859
 RESULT 9
 B1258532 848 bp mRNA linear EST 17-JUL-2001
 LOCUS B1258532

DEFINITION 602972340F1 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:5112032 5',
 mRNA sequence.
 BI258532
 VERSION BI258532.1 GI:14814971
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 848)
 NIH-MGC http://mgc.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished
 REFERENCE Contact: Robert Strausberg, Ph.D.
 TITLE Email: cga@bbs-r@mail.nih.gov
 JOURNAL Tissue Procurement: ATCC
 COMMENT cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: Incyte Genomics, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
 Plate: LHAM1271 row: 0 column: 09
 High quality sequence stop: 839.
 Location/Qualifiers
 1. 848
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5112032"
 /tissue_type="cervical carcinoma cell line"
 /lab_host="DH10B"
 /clone_lib="NIH_MGC_12"
 /note="Organ: cervix; Vector: PCW-SPORT6; Site 1: NCI; Site 2: Salt; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.4 Kb. Library prepared by Life Technologies."
 Technology.

BASE COUNT 196 a 219 c 271 g 162 t
 ORIGIN

Alignment Scores:
 Pred. No.: 3.52e-145 Length: 848
 Score: 1300.50 Matches: 263
 Percent Similarity: 94.29% Conservative: 1
 Best Local Similarity: 93.93% Mismatches: 2
 Query Match: 89.69% Indels: 14
 DB: 12 Gaps: 1

US-09-896-522-2 (1-277) x BI258532 (1-848)

QY 1 MetAlaSerAlaGlyGlyGluAspCysGluSerProAlaProGluAlaAspArgProHis 20
 Db 6 ATGGCTTCGGCGGAGGCGGAAGACTCGCAGAGCCGCCGCGGAGCCGACCTCCAC 65
 QY 21 GlnArgProPheLeuIleGlyValSerGlyGlyThrAlaSerGlyLysSerThrValCys 40
 Db 66 CAGCGGCCCTTCCTGATAGG-----TCGACCGCTGT 98
 QY 41 GlnuylIleMetGluLeuGluGlyGlnAsnGluValGluGlnArgGlnArgLysValVal 60
 Db 99 GAGAAAGTCAATGAGTGTGGGACAGAAACAGAGTGAACGCGGACGCGAAGGTGTGTC 158
 QY 11LeuSerGlnAspArgPheTyrLysValLeuThrAlaGluGlnLysAlaLysAlaLeu 80
 Db 159 ATCTGAGCCAGACGAGTCTTACAAGTCTTGACGCGACAGACGAAAGCCAGGCTTG 218
 QY 81 LysGlyGlnTyrAsnPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThr 100
 Db 219 AAAGGACAGTACATTTTGACATCCAGATGCTTTGATATGATTTGATGACAGACT 278
 QY 101 LeuLysAsnIleValGluGlyLysThrValGluValProThrTyrAspPheValThrHis 120
 Db 279 CTGAAGAACATCGTGGAGGGCAAAACGTGAGGTGCGACCTATGATTTTGTGACACAC 338

QY 121 SerArgLeuProGluThrThrValValTyrProAlaAspValValLeuPheGluGlyIle 140
 Db 339 TCAGGTTACCAAGACACCGGCTGCTTACCTCGGACGCTGTTCTGTTAGGGCATC 398
 QY 141 LeuValPheTyrSerGlnGluIleArgAspMetPheHisLeuArgLeuPheValAspThr 160
 Db 399 TTGGTGTTCACAGCCAGAGATCCGGACATGTTCCACTGCGCTTCTGTCGACACC 458
 QY 161 AspSerAspValArgLeuSerArgValLeuArgAspValArgArgGlyArgAspLeu 180
 Db 459 GACTCCGACATCGAGGCTGCTCCAGAGTTCCTCGGACGCTGGCCGAGGAGGACCTG 518
 QY 181 GlnGlnIleLeuThrGlnTyrThrThrPheValLysProAlaPheGluGluPheCysLeu 200
 Db 519 GAGCAGATTTTGACGACGATACACCACTTGTGAAAGCCGCTTCCAGAGATTTGCTGCTG 578
 QY 201 ProThrLysTyrAlaAspValIleIleProArgGlyValAlaAspAsnMetValAlaIle 220
 Db 579 CCGACAAAGAAATGATGCCGATGTGATCATCCACGAGAGTGACATATGTTGCCATC 638
 QY 221 AsnLeuIleValGlnHisIleGlnAspIleLeuAsn-GlyAspIleCysLysThrPheSar 240
 Db 639 AACCTGATCGTCAGACATCCAGACATTTGAAATTGGACATCTGCAGAAATGCACCG 698
 QY 240 GGLYGLYSerAsnGlyArg--SerTyrLysArgThrPheSerGluProGlyAspHisPro 259
 Db 699 AGAGAGGTCCTCAATGGCGGAGGAGCTTACAGCGGAGACTTTTCTAGCCAGAGGACCACTT 758
 QY 260 GlyMetLeuThrSerGlyLysArg-SerHisLeuGlu-SerSerSerArg 275
 Db 759 GGGATGCTGACTCTGTGGCAAAACGGTGCACATTTGGAATGCCAGACACAG 808

RESULT 10
 BM917506
 LOCUS
 DEFINITION
 AGENCOURT 6606568 NIH_MGC_106 Homo sapiens cDNA clone IMAGE:5484033
 5', mRNA sequence.
 BM917506
 BM917506.1 GI:19367885
 EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 1044)
 NIH-MGC http://mgc.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished
 REFERENCE Contact: Robert Strausberg, Ph.D.
 TITLE Email: cga@bbs-r@mail.nih.gov
 JOURNAL Tissue Procurement: Dr. Daniel McVicar, DBS/NCI
 COMMENT cDNA Library Preparation: Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
 Plate: LNCM2012 row: k column: 10
 High quality sequence stop: 624.
 Location/Qualifiers
 1. 1044
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5484033"
 /tissue_type="natural killer cells, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_106"
 /note="Organ: blood; Vector: pORB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dt priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the

laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH-MGC library."

BASE COUNT 224 a 306 c 318 g 196 t

Alignment Scores:

Pred. No.: 2,42e-142 Length: 1044
Score: 1278.00 Matches: 254
Percent Similarity: 95.56% Conservative: 4
Best Local Similarity: 94.07% Mismatches: 8
Query Match: 88.14% Indels: 4
Gaps: 2

US-09-896-522-2 (1-277) x BW917506 (1-1044)

QY 5 G1VGIYGIUASPCYSGIUSERPROIALPROG1UUA1AASPARPROHISGINARGProphe 24
Db 1 GGAGGCGAAGACTCCGAGAGCCCGCCCGAGGCCACCGTCCGACACCGCCCTTC 60
QY 25 LeuIIeGIYValSerGIYThraISerGIYLySerThraValCysGIUValIleMet 44
Db 61 CTGATAGGGGTGAGCGCGGCACTGCGAGCGGAACTGACCGTGTGAGAAAGATCATG 120
QY 45 GIUleuIleuGIYInaenGIYUaIGIUGInaRGIYnARGIYValIleleuSerGIIn 64
Db 121 GAGTGTGTGAGACGAACGAGGTGAAACAGCGGCGAGGAGGTGATCTCGAGCGAG 180
QY 65 ASPARPhETyrlYsValIleuThraIGIUGInLySaIaLySaIeUlySGIYGIInTyR 84
Db 181 GACAGGTTCNCAAGGCTCTGACGCGACAGACGAAGGCCCTTGAAAGACACTAC 240
QY 85 AsnPhEasphIeSPROaSPaIaPheASPaSnaSPleuNechIaRGTThrLeuLySaenIIe 104
Db 241 AATTTGACCATCGAGATGCTTGTGATATGATTTGATGACAGACTCTGAAGAAATC 300
QY 105 ValIGIUGIYSerThraIGIUGIValProThrThraSPheValThraISerARyLeuPro 124
Db 301 GTGGAGGGCAAAACGGTGAAGGTCCGACCTATGATTTTGGACACACTCAAGGTACCA 360
QY 125 GIUThrThraValValTyRProAlaASpValIValIleuPheGIYIleUValPheTyR 144
Db 361 GAGACCCAGGTGTCTTACCTCGCGAGCTGGTCTGTGAGGCGACTTGGTGTTCAC 420
QY 145 SerGIInuIleARgASpMetPheHISleuRyLeuPheValAspThraSPerASpVal 164
Db 421 AGCCAGAGATCCGGGACATGTTCCACTGGGCTCTTCGTGACACCGACTCGAGCTC 480
QY 165 ArgLeuSerARgARgValIleuARgASpValARgARGIYARgASpLeuGIUGInIleU 184
Db 481 AGGCTGTCTGAAGAGTTCCTCGGAGCGTGGCGGAGGAGGAGCTCGAGACAGATTCTG 540
QY 185 ThrGIInTyRThrThraPheValIySPROAlaPheGIUGIUGIUPheCysLeuPOThrLyS 204
Db 541 ACCGAGTACACCACTTCGTGAAAGCCGCTTCGAGAGATTTCTCTCGCCGACAAAGAG 600
QY 205 TyRAlaASpValIleIleProARGIYValASpASnMetValAlaIleASnLeuIleVal 224
Db 601 TATGCCGATGTGATCATCCACGAGAGTGGACAAATATGTTGCCATCAACTGATCGTG 660
QY 225 G1HIS1IleGIInaSPRIleUaenGIYASpIleCysLySThrPHISARGIYGIYSerASn 244
Db 661 CAGACACATCCAGGACATTCGAATGTGACATCTGCAAAATGGCACCGAAGAGGTCCAT 720
QY 245 G1YARg-SerTyrlYsARg--ThrPhEserGIUPROGIYASpHISPro--GIYMetIle 262
Db 721 GGGGGGAGCTACACAGCGGAGACTTTTCTGAGCCCAAGGGAAACCCCTCGGGAGTCT 780
QY 262 u-ThrSerGIYLySARgSerHISleu 270
Db 781 GAACCTCTGGGCAAAACGTCACCTT 806

RESULT 11
BQ719741 956 bp mRNA linear EST 16-JUL-2002
LOCUS BQ719741
DEFINITION AGENCOUNT_8229663_lupski_dorsal_root_ganglion Homo sapiens cDNA
clone IMAGE:6184625 5', mRNA sequence.

ACCESSION BQ719741
VERSION BQ719741.1 GI:21858638
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 956)
REFERENCE NIH-MGC <http://mgi.nci.nih.gov/>
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
JOURNAL Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Dr. James R. Lupski
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: LLAM13574 row: b column: 18
High quality sequence stop: 552.
Location/Qualifiers

FEATURES

source
1..956
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6184625"
/sex="male"
/tissue_type="dorsal root ganglia"
/dev_stage="adult, 36 yr"
/lab_host="DH10B"
/clone_lib="lupski_dorsal root ganglion"
/note="Vector: PCW-SPORT (Life Technologies); Site_1:
NotI; Site_2: SalI; cDNA made by oligo-dT priming.
Directionally cloned using the following adaptors:
5'-TCGACCCACGCGCCG-3' and
5'-GACTAGTCTGATGCGAGCGGCGCCCT(15)-3'. Size selected >
1 kb for average insert length 1.7 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
College of Medicine) and is available through Life
Technologies."

BASE COUNT 202 a 272 c 296 g 186 t

Alignment Scores:

Pred. No.: 2.79e-142 Length: 956
Score: 1277.00 Matches: 254
Percent Similarity: 95.19% Conservative: 3
Best Local Similarity: 94.07% Mismatches: 10
Query Match: 88.07% Indels: 3
Gaps: 0

US-09-896-522-2 (1-277) x BQ719741 (1-956)

QY 1 MetAlaSerAlaGIYGIUASPCYSGIUSERPROIALPROG1UUA1AASPARPROHIS 20
Db 7 ATGGCTTCGGCGGAGAGCTCTCGAGAGCCCGCCCGGAGGCCACCGTCCGAC 66
QY 21 G1nARgProPheLeuIIeGIYValSerGIYThraISerGIYLySerThraValCys 40
Db 67 CAGCGGCGCTTCTGTATAGGGGTGAGCGCGGCACTGCGAGGAAAGTGCACCTGTGT 126
QY 41 GIUlySIIeMetGIUleuIleuGIYInaenGIYUaIGIUGInaRGIYnARGIYValIle 60
Db 127 GAGAGATCATGAGATTGCTGGGACAGAACGAGGTGAAACAGCGGAGAGGTGTCT 186

Qy 61 ILeuSerGlnAspArgPheTyrLysValLeuThrAlaGluGlnLysAlaLysAlaLeu 80
| | | | |
Db 187 ATCCGAGCCAGGAGAGTTCTACAGAGTCCGAGCGGAGAGGAGAGCCAGGCCCTTG 246
| | | | |
Qy 81 LysGlyGlnTyrAsnPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThr 100
| | | | |
Db 247 AAAGACAGTCAATTTTGACCATCCAGATGCTTTGTAATGATTTATGCACAGGACT 306
| | | | |
Qy 101 LeuLysAsnLysAlaGluGlyLysThrValGluValProThrTyrAspPheValThrHis 120
| | | | |
Db 307 CTGAGAGACATCGTGGAGGGCAAAACGGTGGAGGTGCCAGCCTATGATTTGTGCACAC 366
| | | | |
Qy 121 SerArgLeuProGluThrThrValValTyrProAlaAspValValLeuPheGluGlyTle 140
| | | | |
Db 367 TCAAGTTACCAAGAGACACCGGTGGTCTACCTCGGAGCGTGGTCTCTTTGAGGGCATC 426
| | | | |
Qy 141 LeuValPheTyrSerGlnGlnLysLeuArgMetPheHisLeuArgLeuPheValAspThr 160
| | | | |
Db 427 TTGGTGTCTACAGCAGGAGAGATCCGGACATGTTCCACCTGCCCTCTTGAGACACC 486
| | | | |
Qy 161 AspSerAspValArgLeuSerArgArgValLeuArgAspValArgArgGlyArgAspLeu 180
| | | | |
Db 487 GACTCCGAGCTGAGGCTGTCTGAGAGATCTCCGGAGCGTGGCGGAGGAGGAGACTTG 546
| | | | |
Qy 181 GluGlnLysLeuThrGlnTyrThrThrPheValLysProAlaPheGluGluPheCysLeu 200
| | | | |
Db 547 GAGCAGATTCGTAGCAGCATACACACCTTCGTGAAGCGGCGCTTGAGAGATTCGCTG 606
| | | | |
Qy 201 ProThrLysLysTyrAlaAspValLysLeuProArgGlyValAspAsnMetValAlaI 220
| | | | |
Db 607 CCGACAAAGAGTATGCGAGTGTGATCCCGCAGAGAGTGCACAATATGTTTGCCAT 666
| | | | |
Qy 220 AsnLeuLysValGlnHisLysIleGlnAspIleLeuAsnGlyAspLysCysLysTPHisAr 240
| | | | |
Db 667 CAACCTGATCTGTCAGCAGCATCCAGAGCTTCTTATGGGAACTCTGCAATAGCAGACC 726
| | | | |
Qy 240 gGlyGlySerAsnGlyArgSerTyrLysArgThrPheSerGluPro-GlyAspHisPro 259
| | | | |
Db 727 AGGAAGGTCCAATGGGCGGACCTCCAAAGGAGACCTTTCTTAACCAAGGAGAACCCCT 786
| | | | |
Qy 260 GlyMetLeuThrSerGlyLysArg 267
| | | | |
Db 787 GGGATGTTGACCTTGCCCTAAG 810
| | | | |
RESULT 12
BUS11803 943 bp mRNA linear EST 12-SEP-2002
LOCUS
DEFINITION AGNCOURT_10116119 NIH_MGC_134 Mus musculus cDNA clone
IMAGE:6506624 5', mRNA sequence.
ACCESSION BUS11803
VERSION BUS11803.1 GI:22818036
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 943)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaabs-remail.nih.gov
Tissue Procurement: Dr. David Rowe
CDNA Library Preparation: Invitrogen Corp
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: LHAMJ4069 row: 9 column: 09
High quality sequence stop: 883.
FEATURES
Source 1..943
Location/Qualifiers

/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:6506624"
/tissue_type="undifferentiated limb"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 134"
/note="Vector: pCMV-SPORT6.1, cdb; Site 1: EcoRV; Site 2:
NotI; Cloned unidirectionally. Primer: Oligo dt. Average
insert size 1.7 kb. Constructed by Resgen, Invitrogen
Corp. Note: This is a NIH MGC Library."
BASE COUNT 221 a 263 c 277 g 181 t 1 others
ORIGIN
Alignment Scores:
Pred. No.: 6 25e-142 Length: 943
Score: 1274.00 Matches: 247
Percent Similarity: 94.89% Conservative: 13
Best Local Similarity: 90.15% Mismatches: 12
Query Match: 87.86% Indels: 2
DB: 13 Gaps: 0
US-09-896-522-2 (1-277) x BUS11803 (1-943)
Qy 6 GlyGluAspCysGluSerProAlaProGluAlaAspArgProHisGlnArgProPheLeu 25
| | | | |
Db 3 GGGCGCGGCTCGAGAGCGCGCGCGCGCGAGCCGATGCTCCAGCCGCGCGCTTCCTC 62
| | | | |
Qy 26 ILeGlyValSerGlyGlyThrAlaSerGlyLysSerThrValLysGluLysIleMetGlu 45
| | | | |
Db 63 ATGGGCTGAGCGGCGCGCACCGCTATGCGCAATCAACATGTGTGAGAAATCATAGAG 122
| | | | |
Qy 46 LeuLysGlyGlnAsnGlyValGluGlnArgGlnArgLysValValLysLeuSerGlnAsp 65
| | | | |
Db 123 CTGTGGGACAGAAACGAGAGTGAACCGCGGACGGAAGTGGTCACTCGAGCCAGGAC 182
| | | | |
Qy 66 ArgPheTyrLysValLeuThrAlaGluGlnLysAlaLysAlaLeuLysGlyGlnTyrAsn 85
| | | | |
Db 183 TGCTTCTACAGGTTCTGAGCGCGCAGACAGAGCCCAAGCTTTGAAGGAGACATCAAT 242
| | | | |
Qy 86 PheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThrLeuLysAsnLysVal 105
| | | | |
Db 243 TTGACCAACCAAGATGCTTTGTAATGATCATGCACAAAGCCCTGAAACCAATTGTT 302
| | | | |
Qy 106 GluGlyLysThrValGluValProThrTyrAspPheValThrHisSerArgLeuProGlu 125
| | | | |
Db 303 GAAGGCAAACTGTGAGGTCCCTACTATGATTTGTGACCACTCAAGGTTACAGAG 362
| | | | |
Qy 126 ThrThrValValTyrProAlaAspValValLeuPheGluGlyTleLeuValPheTyrSer 145
| | | | |
Db 363 ACCAGTGTGTCATACCAAGCTGATGTGTCTGTTGAGAGGCAATCTTGATTTCTACACC 422
| | | | |
Qy 146 GlnGluLysLeuArgAspMetPheHisLeuArgLeuPheValAspThrAspSerAspValArg 165
| | | | |
Db 423 CAGAGATCCGGAGACATGTTCCACCTGCCCTCTTTGTGACACACAGACTCTGATGTTAGG 482
| | | | |
Qy 166 LeuSerArgArgValLeuArgAspValArgArgGlyArgAspLeuGlnLysLeuThr 185
| | | | |
Db 483 CTGCTCGAAGAGTCTCCGGGATGTGCAAGGAGAGGAGGAGCTGAGCAGATCCGACT 542
| | | | |
Qy 186 GlnTyrThrThrPheValLysProAlaPheGluGluPheCysLeuProThrLysLysTyr 205
| | | | |
Db 543 CAGTACACCGCCCTTTGTGAAGACAGGCTTTGAGGATTCGCTGCGCATAGGAAGTAC 602
| | | | |
Qy 206 AlaAspValLysLeuProArgGlyValAspAsnMetValAlaLysLeuLysValGln 225
| | | | |
Db 603 GCTGAGGTGTCATCCCTCGAGGAGTGTGATATATGTGGCCATCAACCGATCGTGCAA 662
| | | | |
Qy 226 HisLysGlnAspLysLeuAsnGlyAspLysCysLysTyrPheHisArgGlyGlySerAsnGly 245
| | | | |
Db 663 CACATCCAGGACATCCTNACAGGGGACCTGTGCAAGGAGCCGAGGAGCCCAACGGG 722
| | | | |
Qy 246 ArgSerTyrLysArgThrPheSerGluProGlyAspHisPro-GlyMetLeuThrSerG 265
| | | | |


```
QY 7 GluAspCysGluSerProAlaProGluAlaAspArgProHisGlnArgProPheLeuIle 26
   |||||
Db 2 GAAGACTGGCAGAGACCCCGCGCCGAGAGCCGACCTCCGACACGAGCCCTTCTGATA 61
QY 27 GlyValSerGlyGlyThrAlaSerGlyLysSerThrValCysGluLysIleMetGluLeu 46
   |||||
Db 62 GGGGTGAGCGGCGGCACTGGCCAGCGGGAAGTCGACCGGTGTGTGAAGAATCATGAGAGTTG 121
QY 47 LeuGlyGlnAsnGluValGluGlnArgGlnArgGlyValIleLeuSerGlnAspArg 66
   |||||
Db 122 CTGGGACAGAACGAGCGAACACAGCGGACGGAAGGTGTGCATCTGAGCCAGACAGCAGG 181
QY 67 PheTyrLysValLeuThrAlaGluGlnLysAlaLeuLysGlyGlyTyrAsnPhe 86
   |||||
Db 182 TTCTACAAAGGTCTGACGCGACGACAGACAGCCCAAGGCTTGAAAGACAGATCAATTTT 241
QY 87 AspHisProAspAlaPheAspAsnAspLeuMetHisArgThrLeuLysAsnIleValGlu 106
   |||||
Db 242 GACCATCCAGATGCTTGTGATTAATGATTGATGACAGAGACTCGAAGAACATGTGGAG 301
QY 107 GlyLysThrValGluValProThrTyrAspPheValThrHisSerArgLeuProGluThr 126
   |||||
Db 302 GGCAGAAACGGTGGAGGTGCCGACCTATGATTGTGTGACACACTCAAGTTACAGAGACC 361
QY 127 ThrValValTyrProAlaAspValValLeuPheGluGlyIleLeuValPheTyrSerGln 146
   |||||
Db 362 ACGGTGTCTAACCTCGCGAGCGTGTCTGTGTGAGGACACTGTGTGTCTACAGCCAG 421
QY 147 GluIleArgAspMetPheHisLeuArgLeuPheValAspThrAspSerAspValArgLeu 166
   |||||
Db 422 GAGATCCGGGACATGTTCCACCTGCGCTCTGTGTGACACCGACTCCGAGCTCAGGCTG 481
QY 167 SerArgArgValLeuArgAspValArgArgGlyArgAspLeuGluGlnIleLeuThrGln 186
   |||||
Db 482 TCTCGAAGAGTTCCTCCGCGACGTGCGCGAGGAGGAGCCTGGAGCAGATTCTGACCCAG 541
QY 187 TyrThrThrPheValLysProAlaPheGluGluPheCysLeuProThrLysValTyrAla 206
   |||||
Db 542 TACACCACTTCGTGAGAGCCGAGCTTCGAGAGATTCTGCTGCGACAAAGAGTATGCC 601
QY 207 AspValIleIleProArgGlyVal-AspAsnMetVal-AlaIleAsnLeuIleValGlnH 226
   |||||
Db 602 GATGTATCATCCCGCAGAGGAGTTGACAAATATGTTGGCCATCAAGCTGATGTGCAGC 661
QY 226 IsIleGlnAspIleLeuAsnGlyAspIleCysLys-TrpHisArg-GlyGlySerAsnG 245
   |||||
Db 662 ACATCCAGGACATTCTGAATGTGTACATCTGCAAATTGGCAGCAGAAAGATCCAAATGG 721
QY 245 YArgSerTyrLysArgThrPheSerGluProGlyAspHisProGlyMetLeuThrSerG 265
   |||||
Db 722 GCGGAGCTACAGGCGACTTTTCT-GAGCA-GGAGACCACCCCTGATTTGTAACCTCTGG 779
QY 265 YLysArgSerHisLeuGluSerSerSerArgPro 276
   |||||
Db 780 C---AAGGGACACATTGGAGCCACGACAGACC 810
```

Search completed: November 25, 2003, 10:07:08
Job time : 2563 secs

This Page Blank (uspto)

CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX
SQ Sequence 277 AA;
Query Match 100.0%; Score 277; DB 22; Length 277;
Best Local Similarity 100.0%; Pred. No. 2.6e-256;
Matches 277; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MASAGEDCESPAPEADRPHQRPFLIGVSGGTAGSKSTVCEKIMELLGONEVEQORQKVV 60
DB 1 MASAGEDCESPAPEADRPHQRPFLIGVSGGTAGSKSTVCEKIMELLGONEVEQORQKVV 60
QY 61 ILSODRFYKVLTAQKAKALKGQYNFHPDAPFNDLMHRTLNIVEGKTVVPTDVFVTH 120
DB 61 ILSODRFYKVLTAQKAKALKGQYNFHPDAPFNDLMHRTLNIVEGKTVVPTDVFVTH 120
QY 121 SRLPETTVVYPADVLFEGILVFYSOEIRDMFHLRLFVDTSDVRLSRVLRDVRGRDL 180
DB 121 SRLPETTVVYPADVLFEGILVFYSOEIRDMFHLRLFVDTSDVRLSRVLRDVRGRDL 180
QY 181 EQLITQYTTFFVKPAFEFECFLPTKKYADVILPRGVDMVAINLIVQHIQDILNGDICKMHR 240
DB 181 EQLITQYTTFFVKPAFEFECFLPTKKYADVILPRGVDMVAINLIVQHIQDILNGDICKMHR 240
QY 241 GGSNGRSYKRTFSEPGDHPGMLTSGKSHLESSSRPH 277
DB 241 GGSNGRSYKRTFSEPGDHPGMLTSGKSHLESSSRPH 277
RESULT 2
AAB93941
ID AAB93941 standard; Protein; 277 AA.
XX
AC AAB93941;
XX
DT 26-JUN-2001 (first entry)
DE Human protein sequence SEQ ID NO:13952.
XX
KM Human; primer; detection; diagnosis; antisense therapy; gene therapy.
XX
OS Homo sapiens.
XX
FN EPI074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-0116126.
XX
PR 29-JUL-1999; 99JP-0248036.
XX
PR 27-AUG-1999; 99JP-0300253.
XX
PR 11-JAN-2000; 2000JP-0118776.
XX
PR 02-MAY-2000; 2000JP-0183767.
XX
PR 09-JUN-2000; 2000JP-0241899.
XX
PA (HELI-) HELIX RES INST.
XX
PI Oka T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
XX
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
XX WPI; 2001-318749/34.
XX
XX Primer sets for synthesizing polynucleotides, particularly the 5602
XX full-length cDNAs defined in the specification, and for the detection
XX and/or diagnosis of the abnormality of the proteins encoded by the
XX full-length cDNAs -
PS Claim 8; SEQ ID 13952; 2537bp + CD ROM; English.
CC The present invention describes primer sets for synthesizing 5602
CC full-length cDNAs defined in the specification. Where a primer set

CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.
XX
SQ Sequence 277 AA;
Query Match 100.0%; Score 277; DB 22; Length 277;
Best Local Similarity 100.0%; Pred. No. 2.6e-256;
Matches 277; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MASAGEDCESPAPEADRPHQRPFLIGVSGGTAGSKSTVCEKIMELLGONEVEQORQKVV 60
DB 1 MASAGEDCESPAPEADRPHQRPFLIGVSGGTAGSKSTVCEKIMELLGONEVEQORQKVV 60
QY 61 ILSODRFYKVLTAQKAKALKGQYNFHPDAPFNDLMHRTLNIVEGKTVVPTDVFVTH 120
DB 61 ILSODRFYKVLTAQKAKALKGQYNFHPDAPFNDLMHRTLNIVEGKTVVPTDVFVTH 120
QY 121 SRLPETTVVYPADVLFEGILVFYSOEIRDMFHLRLFVDTSDVRLSRVLRDVRGRDL 180
DB 121 SRLPETTVVYPADVLFEGILVFYSOEIRDMFHLRLFVDTSDVRLSRVLRDVRGRDL 180
QY 181 EQLITQYTTFFVKPAFEFECFLPTKKYADVILPRGVDMVAINLIVQHIQDILNGDICKMHR 240
DB 181 EQLITQYTTFFVKPAFEFECFLPTKKYADVILPRGVDMVAINLIVQHIQDILNGDICKMHR 240
QY 241 GGSNGRSYKRTFSEPGDHPGMLTSGKSHLESSSRPH 277
DB 241 GGSNGRSYKRTFSEPGDHPGMLTSGKSHLESSSRPH 277
RESULT 3
AAB16592
ID AAB16592 standard; Protein; 277 AA.
XX
AC AAB16592;
XX
DT 18-APR-2002 (first entry)
DE Human 57658 protein.
XX
XX Human; uridine kinase-like protein; haematopoietic neoplastic disorder;
XX 57658 protein; leukaemia; haematopoietic disorder; diabetes mellitus;
XX rheumatoid arthritis; psoriasis; dermatitis; osteoporosis; rickets;
XX sarcoma; myocardial infarction; hypertension; atherosclerosis;
XX Alzheimer's disease; Parkinson's disease; Huntington's disease;
XX gene therapy; chromosome mapping; tissue typing; dermatological;
XX cytoskeletal; osteopathic; cardiac; neuroprotective; nootropic;
XX anticonvulsant.
OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX Modified-site 5..10
XX /label= N-myristoylation_site
FT Domain 25..151

FT	/note= "Kinase uridine transferase ATP-binding phosphotribulokinase monophosphokinas precursor PRK cycle phosphopentokkinase domain"
FT	25..231
FT	/note= "Phosphotribulokinase domain; Protein kinase family domain"
FT	27..32
FT	/label= N-myristoylation_site
FT	30..37
FT	/label= ATP/GTP-binding_site
FT	34..36
FT	/label= Protein_kinase-C_phosphorylation_site
FT	38..41
FT	/label= Casein_kinase_II_phosphorylation_site
FT	100..102
FT	/label= Protein_kinase-C_phosphorylation_site
FT	108..115
FT	/label= Tyrosine_kinase_phosphorylation_site
FT	154..277
FT	/note= "Kinase uridine monophosphokinase transferase ATP-binding kinase-like ribonucleoside pyrimidine FIS cDNA domain"
FT	160..163
FT	/label= Casein_kinase_II_phosphorylation_site
FT	167..169
FT	/label= Protein_kinase-C_phosphorylation_site
FT	202..204
FT	/label= Protein_kinase-C_phosphorylation_site
FT	241..246
FT	/label= N-myristoylation_site
FT	247..249
FT	/label= Protein_kinase-C_phosphorylation_site
FT	251..254
FT	/label= Casein_kinase_II_phosphorylation_site
FT	260..265
FT	/label= N-myristoylation_site
FT	264..266
FT	/label= Protein_kinase-C_phosphorylation_site
FT	264..267
FT	/label= Amidation_site
FT	268..271
FT	/label= Casein_kinase_II_phosphorylation_site
FT	273..275
FT	/label= Protein_kinase-C_phosphorylation_site
PX	WO200202761-A2.
PX	10-JUN-2002.
PX	28-JUN-2001; 2001WO-US21063.
PX	30-JUN-2000; 2000US-216503P.
PA	(MILL-) MILLENNIUM PHARM INC.
PI	Glucksmann MA;
DR	WPI; 2002-140091/18.
DR	N-PDB; AAD27186.
PT	New isolated human uridine kinase family polypeptide 57658, useful for treating hematopoietic neoplastic disorders and disorders of neurons, heart and blood vessels _
XX	Claim 4, Fig 1a, 103pp; English.
CC	The patent discloses human uridine kinase-like polypeptides, designated 57658 and polynucleotides encoding such proteins. 57658 DNAs and proteins are useful for developing diagnostic and therapeutic agents for 57658-mediated or related disorders such as haematopoietic neoplastic disorders (e.g. leukaemia), haematopoietic disorders (e.g. psoriasis, dermatitis, diabetes mellitus), rheumatoid arthritis, disorders of bone (e.g. rickets, osteoporosis), heart (e.g. sarcoma, myocardial infarction), blood vessels

Query Match	Similarity	Score	DB 23	Length	277	Best Local	100.0%	Pred.	No. 2.6e-256	Matches	277	Conservative	0	Mismatches	0	Indels	0	Gaps	0
QY	1	MASAGGEDCSPEADPPHPPRLIGVSGGTASGKSTVCKIMELLGQNEVEQRKRV	60																
DB	1	MASAGGEDCSPEADPPHPPRLIGVSGGTASGKSTVCKIMELLGQNEVEQRKRV	60																
QY	61	ILSDRFKVLVTAQKAKALGXQNPDPDAFDDNLHRTKNTVEGKTVVPTYPDFVTH	120																
DB	61	ILSDRFKVLVTAQKAKALGXQNPDPDAFDDNLHRTKNTVEGKTVVPTYPDFVTH	120																
QY	121	SRLEPPTVVYPADVVLFEGLIVFVSOEIRDMFHLRLFPVDTSDVRLSRVRLDVRGRDL	180																
DB	121	SRLEPPTVVYPADVVLFEGLIVFVSOEIRDMFHLRLFPVDTSDVRLSRVRLDVRGRDL	180																
QY	181	EQILTYTTFVKPAFEFECLEPTKKYADVILPRGVDMVAIYLVOHIDILNGDICCKHR	240																
DB	181	EQILTYTTFVKPAFEFECLEPTKKYADVILPRGVDMVAIYLVOHIDILNGDICCKHR	240																
QY	241	GGSNGRSYKRTFSPRDPHGMILTSGRKSHLESSSRPH	277																
DB	241	GGSNGRSYKRTFSPRDPHGMILTSGRKSHLESSSRPH	277																
RESULT 4																			
AAM41288																			
ID	AAM41288	standard; Protein; 296 AA.																	
XX	AAM41288;																		
AC																			
XX																			
DT	22-OCT-2001	(first entry)																	
XX																			
DE		Human polypeptide SEQ ID NO 6219.																	
XX																			
KW		Human; noctropic; immunosuppressant; cytosolic; gene therapy; cancer;																	
KW		peripheral nervous system; neuropathy; central nervous system; CNS;																	
KW		Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;																	
KW		amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;																	
KW		chemokinetic; thrombolytic; drug screening; arthritis; inflammation;																	
KW		leukaemia.																	
XX																			
OS		Homo sapiens.																	
XX																			
PN		WO200153312-A1.																	
XX																			
PD		26-JUL-2001.																	
XX																			
PF		26-DEC-2000# 2000WO-US34263.																	
XX																			
FR		21-JAN-2000; 2000US-0488725.																	
FR		25-APR-2000; 2000US-0552317.																	
FR		09-JUL-2000; 2000US-0598042.		</															

```

XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D,
PI Wang Z, Wehrman T, Xu C, Xue A, Yang Y, Zhang J,
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX WPI: 2001-442253/47.
DR N-PSDB; AA160444.
XX
XX Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
XX
XX Example 2; SEQ ID NO 6219; 10078pp; English.
XX
XX The invention relates to human nucleic acids (AA157798-AA161369) and
CC the encoded polypeptides (AA038642-AA042213) with nucleotide,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX
XX Sequence 296 AA;
SQ
Query Match 100.0%; Score 277; DB 22; Length 296;
Best Local Similarity 100.0%; Pred. No. 2,8e-256;
Matches 277; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 MASAGDECEPAPADRPFRPFLIGVSGTASGKSTVCEKIMELLGONEVEQRORRV 60
DB MASAGDECEPAPADRPFRPFLIGVSGTASGKSTVCEKIMELLGONEVEQRORRV 79
QY 61 ILSQDRFYKVLTAEOAKALKGQYNFDHPDAFNDLMHRTLKNIVEGTVVEPTVDFVTH 120
DB ILSQDRFYKVLTAEOAKALKGQYNFDHPDAFNDLMHRTLKNIVEGTVVEPTVDFVTH 139
QY 121 SRLPETTVVPADVLFEGILVFYSQEIRDMFHLRLFVDTSDVRLSRVLRDVRGRDL 180
DB SRLPETTVVPADVLFEGILVFYSQEIRDMFHLRLFVDTSDVRLSRVLRDVRGRDL 199
QY 181 EQILTYTTFVKPAFEFECPLPTKKYADVILIPGVDMVAIINLVQHIQDILNGDICKMHR 240
DB EQILTYTTFVKPAFEFECPLPTKKYADVILIPGVDMVAIINLVQHIQDILNGDICKMHR 259
QY 241 GGSNGRSYKRTFSEPDHPGMLTSGKRSHTLESSSRPH 277
DB GGSNGRSYKRTFSEPDHPGMLTSGKRSHTLESSSRPH 296
XX
XX RESULT 5
XX AA014412
XX ID AA014412 standard; Protein; 260 AA.
XX
XX AA014412;
XX
XX 02-MAY-2002 (first entry)
XX
XX Protein of a human uridine kinase (UDK).
XX
XX Human: uridine kinase; diagnostic assay; mutation detection; UDK;
XX probe; chromosome localisation study; tissue expression; gene therapy;
XX antibody; vaccine; human ovarian cancer; immunological disorder;
XX human colon carcinoma; immunogen.
XX
XX Homo sapiens.
XX
OS

```

```

XX
XX NO200172963-A2.
XX
XX 04-OCT-2001.
XX
XX 27-MAR-2001; 2001WO-US09663.
XX
XX 27-MAR-2000; 2000US-0536647.
XX
XX (SMIK) SMITHKLINE BEECHAM CORP.
XX
XX Ho YS, Johnson RK;
XX
XX WPI: 2001-626259/72.
XX
XX N-PSDB; AA038735.
XX
XX Novel human uridine kinase polypeptides useful for treating cancers,
PT and to identify agonists and antagonists of the polypeptide useful for
PT treating conditions associated with uridine kinase imbalance -
XX
XX Claim 3; Page 29-30; 31pp; English.
XX
XX The invention relates to newly identified human uridine kinase (UDK)
CC polypeptides and polynucleotides and methods for producing such
CC polypeptides by recombinant techniques. Also disclosed in the invention
CC are methods for utilising uridine kinase polypeptides and polynucleotides
CC in diagnostic assays. The polynucleotides and polypeptides of the
CC invention may be used as diagnostic reagents by detecting mutations in an
CC associated gene. An array of oligonucleotide probes comprising the
CC uridine kinase polynucleotide sequence or fragments thereof can be
CC constructed to conduct efficient screening of genetic mutations, for
CC example, detection of abnormally decreased or increased levels of
CC polypeptide or mRNA expression may also be used for diagnosing or
CC determining susceptibility of a subject to a disease of the invention.
CC The polynucleotide sequences of the invention can be used for chromosome
CC localisation studies and tissue expression studies. The polypeptides of
CC the invention or fragments thereof may be used as immunogens to produce
CC antibodies. These antibodies may be employed to isolate or identify
CC clones expressing the polypeptide. The polypeptides and polynucleotides
CC of the invention can be used as a vaccine or in gene therapy to treat
CC diseases such as human ovarian cancer, human colon carcinomas, and
CC immunological disorders. This sequence represents the protein of a human
CC uridine kinase of the invention.
XX
XX Sequence 260 AA;
SQ
Query Match 93.9%; Score 260; DB 22; Length 260;
Best Local Similarity 100.0%; Pred. No. 4,4e-240;
Matches 260; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 MASAGDECEPAPADRPFRPFLIGVSGTASGKSTVCEKIMELLGONEVEQRORRV 60
DB MASAGDECEPAPADRPFRPFLIGVSGTASGKSTVCEKIMELLGONEVEQRORRV 60
QY 61 ILSQDRFYKVLTAEOAKALKGQYNFDHPDAFNDLMHRTLKNIVEGTVVEPTVDFVTH 120
DB ILSQDRFYKVLTAEOAKALKGQYNFDHPDAFNDLMHRTLKNIVEGTVVEPTVDFVTH 120
QY 121 SRLPETTVVPADVLFEGILVFYSQEIRDMFHLRLFVDTSDVRLSRVLRDVRGRDL 180
DB SRLPETTVVPADVLFEGILVFYSQEIRDMFHLRLFVDTSDVRLSRVLRDVRGRDL 180
QY 181 EQILTYTTFVKPAFEFECPLPTKKYADVILIPGVDMVAIINLVQHIQDILNGDICKMHR 240
DB EQILTYTTFVKPAFEFECPLPTKKYADVILIPGVDMVAIINLVQHIQDILNGDICKMHR 240
QY 241 GGSNGRSYKRTFSEPDHPG 260
DB GGSNGRSYKRTFSEPDHPG 260
XX
XX RESULT 6
XX AA064506

```

ID AAG64506 standard; Protein; 276 AA.
XX
AC AAG64506;
XX
DT 02-OCT-2001 (first entry)
XX
DE Human uridine kinase.
XX
KW Human; uridine kinase; UK.
XX
OS Homo sapiens.
XX
PN CN1287172-A.
XX
PD 14-MAR-2001.
XX
PF 07-SEP-1999; 99CN-0118818.
XX
PR 07-SEP-1999; 99CN-0118818.
XX
PA (UYFU-) UNIV FUDAN.
XX
PI Yu L, Zhao Y, Zhang H;
XX
DR WPI; 2001-409529/44.
DR N-PSDB; AAH75355.
XX
PT Human uridine kinase and its coding sequence, preparation and
XX application -
XX
PS Claim 2; Page 15-16 (Disclosure); 20pp; Chinese.
XX
CC The invention relates to human uridine kinase (UK).
XX
SQ Sequence 276 AA;
XX
Query Match 68.2%; Score 189; DB 22; Length 276;
Best Local Similarity 100.0%; Pred. No. 3.3e-172;
Matches 189; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 58 KVVILSQDRFYKVLTAEOKAKALKGQYNFDPDAFNDLMHRTLKNIVEGKTYEVPPTYDF 117
DB 57 KVVILSQDRFYKVLTAEOKAKALKGQYNFDPDAFNDLMHRTLKNIVEGKTYEVPPTYDF 116
QY 118 VTHSRLEPETYVVPADVLPFGILVFSQETRDMFHLRLFYDTSVRLSRVLRDVRG 177
DB 117 VTHSRLEPETYVVPADVLPFGILVFSQETRDMFHLRLFYDTSVRLSRVLRDVRG 176
QY 178 RDLFQILTOYTFPKPAFEFCLPTKXKADVIIPRGVNMVAIYLVOHIDILNGDICK 237
DB 177 RDLFQILTOYTFPKPAFEFCLPTKXKADVIIPRGVNMVAIYLVOHIDILNGDICK 236
QY 238 WHRGSGNGR 246
DB 237 WHRGSGNGR 245
XX
RESULT 7
AB889353
ID AB889353 standard; Protein; 190 AA.
XX
AC AB889353;
XX
DT 24-MAY-2002 (first entry)
XX
DE Human polypeptide SEQ ID NO 1729.
XX
KW Cytosolic; immunosuppressive; nootropic; neuroprotective; antiviral;
XX antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor;
XX vulnerable; anticonvulsant; antibacterial; antifungal; antiparasitic;
XX cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
XX neurological disease; infection; human; secreted protein.
XX

OS Homo sapiens.
XX
PN NO200190304-A2.
XX
PD 29-NOV-2001.
XX
PF 18-MAY-2001; 2001WO-US16450.
XX
PR 19-MAY-2000; 2000US-205515P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Birse CE, Rosen CA;
XX
DR WPI; 2002-122018/16.
DR N-PSDB; ABL89762.
XX
PT Novel 1405 isolated polypeptides, useful for diagnosis, treatment and
PT prevention of neural, immune system, muscular, reproductive,
PT gastrointestinal, pulmonary, cardiovascular, renal and proliferative
PT disorders -
XX
PS Claim 11; SEQ ID NO 1729; 2081pp + Sequence Listing; English.
XX
CC The invention relates to novel genes (ABL89449-ABL90853) and proteins
CC (AB889040-AB890444) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful
CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
CC marrow, breast, gastrointestinal tract, liver, lung or urogenital;
CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune;
CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
CC colitis; (c) cardiovascular disorders such as myocardial ischaemia;
CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
CC and parasitic infections.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at fcp.wipo.int/pub/publ/published_pct_sequences.
XX
SQ Sequence 190 AA;
XX
Query Match 49.1%; Score 136; DB 23; Length 190;
Best Local Similarity 100.0%; Pred. No. 1.1e-121;
Matches 136; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 34 SGKSTYCEKIMELGONEVEORQRKVILISORFVVLTAEOKAKALKGQYNFDPDAF 93
DB 34 SGKSTYCEKIMELGONEVEORQRKVILISORFVVLTAEOKAKALKGQYNFDPDAF 93
QY 94 NDLMRRTLKNIVEGKTYEVPDYVTHSRLEPETYVVPADVLPFGILVFSQETRDMFHL 153
DB 94 NDLMRRTLKNIVEGKTYEVPDYVTHSRLEPETYVVPADVLPFGILVFSQETRDMFHL 153
QY 154 LRLFVDTSDVRLSRR 169
DB 154 LRLFVDTSDVRLSRR 169
XX
RESULT 8
ABG70328
ID ABG70328 standard; Protein; 120 AA.
XX
AC ABG70328;
XX
DT 21-OCT-2002 (first entry)
XX
DE Human MMDT protein Incyte ID No. LI:235557.12.orf2:2001JAN12.
XX
KW Human; molecule for disease detection and treatment; MMDT; cancer;
XX

KW cell proliferative disorder; arteriosclerosis; cirrhosis; asthma;
 KW autoimmune disorder; inflammatory disorder; Crohn's disease;
 KW multiple sclerosis; cystostatic; antiarteriosclerotic; antiinflammatory;
 KW hepatotropic; immunosuppressive; antiasthmatic.
 XX Homo sapiens.
 XX WO200255738-A2.
 XX
 XX 18-JUL-2002.
 XX
 XX 09-JAN-2002; 2002WO-US01008.
 XX
 XX 12-JAN-2001; 2001US-261622P.
 XX 16-JAN-2001; 2001US-261865P.
 XX 17-JAN-2001; 2001US-262208P.
 XX 17-JAN-2001; 2001US-262209P.
 XX 17-JAN-2001; 2001US-262326P.
 XX 19-JAN-2001; 2001US-263063P.
 XX 19-JAN-2001; 2001US-263065P.
 XX 19-JAN-2001; 2001US-263329P.
 XX
 XX (INCYTE GENOMICS INC.
 XX Panzer SR, Lincoln SE, Altus CM, Dufour GE, Hillman JL, Jones AL;
 XX Dam TC, Liu TP, Harrie B, Flores V, Daffo A, Marwaha R, Chen AJ;
 XX Chang SC, Gerstein EH, Peralta CH, David MH, Lewis SA;
 XX WPI; 2002-590679/63.
 XX N-PSDB; ABS51801.
 XX
 XX New disease detection and treatment molecule (MDPT) polynucleotides and
 XX polypeptides, useful in diagnosing, studying, preventing or treating
 XX diseases associated with MDT expression, e.g. autoimmune or
 XX inflammatory disorders -
 XX
 XX Claim 27; Page 120; 123pp; English.
 XX
 XX The present invention relates to the isolation of novel human
 XX molecules for disease detection and treatment (MDPT), and the
 XX polynucleotide sequences (mdpt) encoding them. The MDT polypeptides
 XX may be used to screen for molecules that bind to, or are bound by the
 XX encoded polypeptides, and to develop a transcript image of a tissue or
 XX cell type. Probes comprising at least 20 nucleotides of the mdpt
 XX polynucleotide may be used to assess the toxicity of a test compound.
 XX The MDT polypeptides and mdpt polynucleotides are useful in the
 XX diagnosis, study, prevention and treatment of diseases associated with
 XX the expression of molecules for disease detection and treatment. Such
 XX disorders include cell proliferative disorders (e.g. arteriosclerosis,
 XX cirrhosis, or cancers), and autoimmune/inflammatory disorders
 XX (e.g. asthma, Crohn's disease, or multiple sclerosis). The mdpt
 XX polynucleotides may also be used as molecule markers, in microarrays,
 XX and in somatic or germline gene therapy. ABS70306-ABS70341 represent
 XX the MDT proteins of the invention.
 XX
 XX Sequence 120 AA;
 XX
 XX Query Match 40.4%; Score 112; DB 23; Length 120;
 XX Best Local Similarity 100.0%; Pred. No. 6.4e-99;
 XX Matches 112; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 XX QY 90 DAFNDNDLHRTLNKNIIVEGKTVVPTVDFTHSRLPETTVVYPADVLFEGILVFSQSEIR 149
 XX DAFNDNDLHRTLNKNIIVEGKTVVPTVDFTHSRLPETTVVYPADVLFEGILVFSQSEIR 64
 XX
 XX QY 150 DMFHRLFLVDDSDVRLSRVLRDVRGRDLLEQLITQTTTTPKAFEBFCIP 201
 XX DMFHRLFLVDDSDVRLSRVLRDVRGRDLLEQLITQTTTTPKAFEBFCIP 116
 XX
 XX RESULT 9
 XX AAO14413
 XX ID AAO14413 standard; Protein; 277 AA.

XX
 XX AAO14413;
 XX
 XX 02-MAY-2002 (first entry)
 XX
 XX Protein relating to a human uridine kinase (UDK) of the invention.
 XX
 XX Human; uridine kinase; diagnostic assay; mutation detection; UDK;
 XX probe; chromosome localisation study; tissue expression; gene therapy;
 XX antibody; vaccine; human ovarian cancer; immunological disorder;
 XX human colon carcinoma; immunogen.
 XX
 XX Unidentified.
 XX
 XX WO200172963-A2.
 XX
 XX 04-OCT-2001.
 XX
 XX 27-MAR-2001; 2001WO-US09663.
 XX
 XX 27-MAR-2000; 2000US-0536647.
 XX
 XX (SMITK) SMITHKLINE BEECHAM CORP.
 XX
 XX Ho YS, Johnson RK;
 XX WPI; 2001-626259/72.
 XX
 XX Novel human uridine kinase polypeptides useful for treating cancers,
 XX and to identify agonists and antagonists of the polypeptide useful for
 XX treating conditions associated with uridine kinase imbalance -
 XX
 XX Disclosure; Page 23; 31pp; English.
 XX
 XX The invention relates to newly identified human uridine kinase (UDK)
 XX polypeptides and polynucleotides and methods for producing such
 XX polypeptides by recombinant techniques. Also disclosed in the invention
 XX are methods for utilising uridine kinase polypeptides and polynucleotides
 XX in diagnostic assays. The polynucleotides and polypeptides of the
 XX invention may be used as diagnostic reagents by detecting mutations in an
 XX associated gene. An array of oligonucleotide probes comprising the
 XX uridine kinase polynucleotide sequence or fragments thereof can be
 XX constructed to conduct efficient screening of genetic mutations, for
 XX example. Detection of abnormally decreased or increased levels of
 XX polypeptide or mRNA expression may also be used for diagnosing or
 XX determining susceptibility of a subject to a disease of the invention.
 XX The polynucleotide sequences of the invention can be used for chromosome
 XX localisation studies and tissue expression studies. The polypeptides of
 XX the invention or fragments thereof may be used as immunogens to produce
 XX antibodies. These antibodies may be employed to isolate or identify
 XX clones expressing the polypeptide. The polypeptides and polynucleotides
 XX of the invention can be used as a vaccine or in gene therapy to treat
 XX diseases such as human ovarian cancer, human colon carcinomas, and
 XX immunological disorders. This sequence represents the protein relating to
 XX a human uridine kinase (UDK) of the invention.
 XX NOTE: The present sequence is stated as being the same as that shown as
 XX SEQ ID NO: 2 in the sequence listing of the specification. However, the
 XX sequences differ.
 XX
 XX Sequence 277 AA;
 XX
 XX Query Match 16.2%; Score 45; DB 22; Length 277;
 XX Best Local Similarity 100.0%; Pred. No. 1.4e-34;
 XX Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 XX QY 100 TLKNIIVEGKTVVPTVDFTHSRLPETTVVYPADVLFEGILVY 144
 XX TLKNIIVEGKTVVPTVDFTHSRLPETTVVYPADVLFEGILVY 144
 XX
 XX RESULT 10
 XX AAM38694
 XX ID AAM38694 standard; Protein; 261 AA.

XX AAM38694;
AC
XX
XX 22-OCT-2001 (first entry)
XX
XX Human polypeptide SEQ ID NO 1839.

Human, nocutropic; immunosuppressant; cytostatic; gene therapy; cancer;
KM peripheral nervous system; neuropathy; central nervous system; CNS;
KM Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KM amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KM chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KM leukaemia.

XX Homo sapiens.
XX
XX WO200153312-A1.
XX
XX 26-JUL-2001.

XX 26-DEC-2000; 2000MO-US34263.

XX 21-JAN-2000; 2000US-0488725.
XX 25-APR-2000; 2000US-0552317.
XX 09-JUL-2000; 2000US-0598042.
XX 19-JUL-2000; 2000US-0620312.
XX 03-AUG-2000; 2000US-0653450.
XX 14-SEP-2000; 2000US-0662191.
XX 19-OCT-2000; 2000US-0693036.
XX 29-NOV-2000; 2000US-0727344.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D,
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
XX WPI: 2001-442253/47.
XX N-PSDB; AA157850.

XX Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
XX

XX Example 3; SEQ ID NO 1839; 10078bp; English.

XX The invention relates to human nucleic acids (AA157798-AA161369) and
XX the encoded polypeptides (AAM38642-AAM42213) with nocutropic,
XX immunosuppressant and cytostatic activity. The polynucleotides are useful
XX in gene therapy. A composition containing a polypeptide or polynucleotide
XX of the invention may be used to treat diseases of the peripheral nervous
XX system, such as peripheral nervous injuries, peripheral neuropathy and
XX localised neuropathies and central nervous system diseases, such as
XX Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
XX lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
XX utilisation of the activities such as: Immune system suppression,
XX Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
XX and thrombolytic activity, cancer diagnosis and therapy, drug screening,
XX assays for receptor activity, arthritis and inflammation, leukaemia and
XX C.N.S disorders.
XX Note: The sequence data for this patent did not form part of the printed
XX specification.

XX Sequence 261 AA;

Query Match 9.0%; Score 25; DB 22; Length 261;
Best Local Similarity 100.0%; Pred. No. 1.8e-15;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 189 TFWKPAFEFECPLTKKADVIIRG 213
DB 187 TFWKPAFEFECPLTKKADVIIRG 211

RESULT 11

AA873494
ID AAB73494 standard; Protein; 261 AA.

AC AAB73494;
XX
XX 31-JUL-2001 (first entry)

XX Human transferase HTFS-1, SEQ ID NO:1.

XX Human transferase; HTFS; agonist; antagonist; cellular signalling;
KM proliferation; cell proliferative disorder; immune disorder;
KM atherosclerosis; hepatitis; psoriasis; cancer; tumour;
KM inflammation; AIDS; Addison's disease; allergy; asthma; anaemia;
KM cirrhosis; Crohn's disease; atopic dermatitis; diabetes mellitus;
KM multiple sclerosis; rheumatoid arthritis; pancreatitis;
KM systemic lupus erythematosus; thrombocytopenia; ulcerative colitis;
KM haemodialysis; extracorporeal circulation; trauma; transgenic animal;
KM gene therapy; drug screening.

XX Homo sapiens.

XX WO200132888-A2.

XX 10-MAY-2001.

XX 02-NOV-2000; 2000MO-US30485.

XX 04-NOV-1999; 99US-0163595.

XX (INCY-) INCYTE GENOMICS INC.

XX Tang YT, Yue H, Hillman JL, Lal P, Bandman O, Patterson C;
PI Shih LL, Azimzai Y, Lu DAM, Baughn MR;
XX
XX WPI: 2001-328796/34.
XX N-PSDB; AAH23801.

XX Human transferase polypeptides and polynucleotides useful for
PT diagnosis, prevention and treatment of cell proliferative and immune
PT system disorders and for identifying agonists and antagonists -
XX

XX Claim 1; Page 103-104; 157bp; English.

XX Sequences AAB73494-AAB73535 represent novel human transferase proteins
XX HTFS-1 to HTFS-42, and sequences AAH23801-AAH23842 represent cDNAs
XX encoding them. The proteins play important roles in the regulation of
XX cellular signalling and proliferation. The HTFS proteins are useful for
XX screening compounds for their effectiveness as agonists or antagonists of
XX transferase activity, or for compounds that specifically bind to an HTFS
XX protein or which modulates the activity of an HTFS protein.
XX Pharmaceutical compositions comprising an HTFS protein, HTFS
XX agonist or antagonist, or genetic construct encoding an HTFS
XX protein are useful for treating a disease or condition associated
XX with decreased or increased expression of functional HTFS. Disorders
XX which may be treated using such compositions include cell proliferative
XX disorders and immune disorders. For example, diseases which may be
XX treated include atherosclerosis, hepatitis, psoriasis, cancers (including
XX breast, bladder, bone marrow, brain and uterus cancer), inflammation,
XX AIDS, Addison's disease, allergies, asthma, anaemia, cirrhosis, Crohn's
XX disease, atopic dermatitis, diabetes mellitus, multiple sclerosis,
XX rheumatoid arthritis, pancreatitis, systemic lupus erythematosus,
XX thrombocytopenia, and ulcerative colitis. They may also be used to treat
XX complications of cancer, haemodialysis, extracorporeal circulation,
XX trauma and haematopoietic cancer, including lymphoma, leukaemia and
XX myeloma. Polynucleotides encoding HTFS proteins are useful for creating
XX transgenic animals to model human diseases, for diagnostic purposes and
XX to generate hybridisation probes useful in mapping the naturally
XX occurring genomic sequences. HTFS, and its catalytic or immunogenic
XX fragments are useful for screening libraries of compounds in a variety of
XX drug screening techniques. Antibodies which specifically bind HTFS may be
XX used for the diagnosis of disorders associated with the expression of
XX HTFS, or in assays to monitor patients being treated with HTFS or

CC agonists, antagonists or inhibitors of HTFS. The present sequence
CC represents an HTFS protein of the invention.

Sequence 261 AA;

Query Match 9.0%; Score 25; DB 22; Length 261;
Best Local Similarity 100.0%; Pred. No. 1.8e-15;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 189 TFWKPAFEFCLPTKKYADVILPRG 213
Db 187 TFWKPAFEFCLPTKKYADVILPRG 211

RESULT 12

AAB56582
ID AAB56582 standard; Protein; 337 AA.

AC AAB56582;

DT 13-MAR-2001 (first entry)

DE Human prostate cancer antigen protein sequence SEQ ID NO:1160.

XX Human; prostate cancer; prostate cancer antigen; detection; diagnosis;
XX neuroprotective; cytoskeletal; cardioactive; immunomodulatory; muscular;
XX vulnereary; gastrointestinal; nephrotoxic; antiinfective; gynaecological;
XX antibacterial; gene therapy; neural; immune; reproductive; renal;
XX gastrointestinal; pulmonary; cardiovascular; proliferative disorder;
XX wound; infectious disease.

OS Homo sapiens.

XX WO200055174-A1.

XX 21-SEP-2000.

XX 08-MAR-2000; 2000WO-US05988.

XX 12-MAR-1999; 99US-0124270.

XX (HUMA-) HUMAN GENOME SCI INC.

XX (ROSE/) ROSEN C A.

XX Rosen CA, Ruben SM;

XX WPI; 2000-587513/55.

XX N-PSDB; AAF15785.

PT Prostate cancer associated gene sequences, referred to as prostate
PT cancer antigens, useful for treatment, prevention, and diagnosis of
PT disorders such as prostate cancer -

PS Claim 11; Page 1566-1567; 2338pp; English.

XX AAF15566 to AAF15505 encode the human prostate cancer associated
XX proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.

CC The prostate cancer antigens can have neuroprotective, cytoskeletal,
CC cardioactive, immunomodulatory, muscular, vulnereary, gastrointestinal,
CC nephrotoxic, antiinfective, gynaecological and antibacterial activities,
CC and can be used in gene therapy. The prostate cancer antigen
CC polynucleotides may be used for detection of prostate cancer, chromosome
CC identification, as chromosome markers, and for numerous other diagnostic
CC or research purposes. The prostate cancer antigens may be used to treat
CC disorders such as neural, immune, muscular, reproductive,
CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
CC disorders, wounds, and infectious diseases. AAF15506 to AAF15514 to
CC AAB57303 represent sequences used in the exemplification of the present
CC invention.

Sequence 337 AA;

Query Match 9.0%; Score 25; DB 21; Length 337;

Best Local Similarity 100.0%; Pred. No. 2.2e-15;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 189 TFWKPAFEFCLPTKKYADVILPRG 213
Db 263 TFWKPAFEFCLPTKKYADVILPRG 287

RESULT 13

ABP41393
ID ABP41393 standard; Protein; 337 AA.

AC ABP41393;

DT 22-AUG-2002 (first entry)

DE Human ovarian antigen HOELP29, SEQ ID NO:2525.

XX Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;
XX ovarian cancer; breast cancer; tumour; reproductive system disorder;
XX infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
XX PCOS; ovarian cyst; dysmenorrhoea; endocrine disorder; infection;
XX inflammatory condition; immune disorder; blood disorder;
XX cardiovascular disorder; respiratory disorder; neurological disorder;
XX gastrointestinal disorder; urinary system disorder; drug screening;
XX gene therapy; chromosome mapping; forensic analysis;
XX antibody preparation; cytoskeletal; immunomodulatory; neuroprotective;
XX antiinflammatory; gynaecological; reproductive; chromosome 1p32.

OS Homo sapiens.

XX WO200200677-A1.

XX 03-JAN-2002.

XX 07-JUN-2001; 2001WO-US18569.

XX 07-JUN-2000; 2000US-209467P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Birse CE, Rosen CA;

XX WPI; 2002-147878/19.

XX N-PSDB; ABQ54470.

PT Isolated nucleic acid molecules encoding novel ovarian polypeptides,
PT useful in the prevention, treatment and diagnosis of cancer (e.g.
PT ovarian cancer), immune disorders, cardiovascular disorders and
PT neurological diseases -

PS Claim 11; SEQ ID NO 2525; 2922pp; English.

XX The invention relates to 2175 novel human ovarian antigens (ABP41054-
XX ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also
XX encompasses polypeptides 90% identical and polynucleotides 95% identical
XX to the sequences of the invention. The invention additionally relates to
XX recombinant vectors and host cells comprising human ovarian antigen
XX polynucleotides, antibodies against human ovarian antigens, and the use
XX of ovarian antigen polynucleotides and polypeptides in diagnosing,
XX treating, prognosing or preventing various ovary and/or breast-related
XX disorders. Such conditions include ovarian cancer and breast cancer, and
XX metastatic tumours of ovarian or breast origin, reproductive system
XX disorders (e.g., infertility, disorders of pregnancy, anovulation,
XX polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine
XX disorders, infections (e.g., chlamydia, HIV, toxoplasmosis), and toxic
XX shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
XX vaginitis), immune disorders (e.g., congenital and acquired
XX immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
XX blood-related disorders (e.g., anaemia), cardiovascular disorders,
XX respiratory disorders, neurological disorders, gastrointestinal disorders
XX and urinary system disorders. Ovarian antigen polypeptides and
XX polynucleotides may also be used in screening for compounds which

CC modulate ovarian antigen expression or activity. The polynucleotides may
 CC further be used for gene therapy, chromosome mapping, in the
 CC identification of individuals and in forensic analysis, and the
 CC polypeptides may be used as food additives or to prepare antibodies
 CC useful in disease diagnosis, drug targeting and phenotyping. The present
 CC sequence represents a human ovarian antigen of the invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 337 AA;

Query Match 9.0%; Score 25; DB 23; Length 337;
 Best Local Similarity 100.0%; Pred. No. 2,2e-15;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 189 TPVKPAFEERFCLPTKXYADVIIPRG 213
 |||||
 DB 263 TPVKPAFEERFCLPTKXYADVIIPRG 287

RESULT 14

AAE16594 ID AAE16594 standard; Protein; 125 AA.

XX AAE16594;

DT 18-APR-2002 (first entry)

XX Human 57658 protein prodomain consensus sequence #1.

XX Human; uridine kinase-like protein; haematopoietic neoplastic disorder;
 KM 57658 protein; leukaemia; haematopoietic disorder; diabetes mellitus;
 KM rheumatoid arthritis; psoriasis; dermatitis; osteoporosis; rickets;
 KM sarcoma; myocardial infarction; hypertension; atherosclerosis;
 KM Alzheimer's disease; Parkinson's disease; Huntington's disease;
 KM gene therapy; chromosome mapping; tissue typing; dermatological;
 KM cytoskeletal; osteopathic; cardiac; neuroprotective; nocrotic;
 KM anticonvulsant; prodomain.

OS Homo sapiens.

PN MO200202761-A2.

XX 10-JAN-2002.

PF 28-JUN-2001; 2001WO-US21063.

XX 30-JUN-2000; 2000US-216503P.

XX (MILL-) MILLENNIUM PHARM INC.

PI Glucksmann MA;

DR WPI; 2002-140091/18.

XX New isolated human uridine kinase family polypeptide 57658, useful for
 PT treating hematopoietic neoplastic disorders and disorders of neurons,
 PT heart and blood vessels -

XX Disclosure; Fig 4; 103pp; English.

XX The patent discloses human uridine kinase-like polypeptides, designated
 CC 57658 and polynucleotides encoding such proteins. 57658 DNAs and proteins
 CC are useful for developing diagnostic and therapeutic agents for 57658-
 CC mediated or related disorders such as haematopoietic neoplastic disorders
 CC (e.g. leukaemia), haematopoietic disorders (e.g. psoriasis, dermatitis,
 CC diabetes mellitus, rheumatoid arthritis, disorders of bone (e.g. rickets,
 CC osteoporosis), heart (e.g. sarcoma, myocardial infarction), blood vessels
 CC (e.g. hypertension, atherosclerosis) and neurons (e.g. Huntington's
 CC disease, Alzheimer's disease, Parkinson's disease). 57658 sequences or
 CC their antibodies are useful in screening assays, detection assays (e.g.
 CC forensic biology) and predictive medicine (e.g. diagnostic assays).

CC prognostic assays, and monitoring clinical trials and pharmacogenomics).
 CC They are useful as reagents for diagnosing and treating 57658-mediated
 CC disorders. 57658 DNAs are useful in gene therapy, in chromosome mapping,
 CC to identify an individual from a minute biological sample (tissue typing)
 CC and to aid in forensic identification of the biological sample. The
 CC present sequence is the consensus sequence of the biological sample.
 CC prodomain, kinase uridine monophosphokinase transferase ATP-binding
 CC kinase-like ribonucleoside pyrimidine FIS cDNA domain.

SQ Sequence 125 AA;

Query Match 5.4%; Score 15; DB 23; Length 125;
 Best Local Similarity 100.0%; Pred. No. 3.4e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 246 RSYKRTFSEPDHPG 260
 |||||
 DB 94 RSYKRTFSEPDHPG 108

RESULT 15

ABBE2307 ID ABBE2307 standard; Protein; 260 AA.

XX ABBE2307;

DT 26-MAR-2002 (first entry)

XX Drosophila melanogaster polypeptide SEQ ID NO 13713.

KM Drosophila; developmental biology; cell signalling; insecticide;
 KM pharmaceutical.

OS Drosophila melanogaster.

XX MO200171042-A2.

XX 27-SEP-2001.

PF 23-MAR-2001; 2001WO-US09231.

XX 23-MAR-2000; 2000US-191637P.

PR 11-JUL-2000; 2000US-0614150.

PA (PEKE) BE CORP NY.

PI Venter JC, Adams M, Li PMD, Myers EW;

DR WPI; 2001-656860/75.

DR N-PSDB; ABL06410.

XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -

XX Disclosure; SEQ ID NO 13713; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABU16176-ABU30511), expressed DNA
 CC sequences (ABBS7737-ABBS7072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 260 AA;

Query Match 5.4%; Score 15; DB 22; Length 260;
 Best Local Similarity 100.0%; Pred. No. 6.4e-06;

	Matches	15;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
Oy	130		YPADVLFEGILVFY	144						
Db	135		YPADVLFEGILVFY	149						

Search completed: November 25, 2003, 08:09:03
Job time : 39 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2003, 08:08:15 ; Search time 19 Seconds
(without alignments)
616.848 Million cell updates/sec

Title: US-09-896-522-2
Perfect score: 277
Sequence: 1 MASAGEDCESPAPEADRP...HPGLTRGKRSHLESSSRPH 277

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 328717 seqs, 42310858 residues

Word size : 15

Total number of hits satisfying chosen parameters: 2

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : Issued Patents_AA:*
1: /cgn2_6/ptodata/2/1aa/5A COMB .pep:*
2: /cgn2_6/ptodata/2/1aa/5B COMB .pep:*
3: /cgn2_6/ptodata/2/1aa/6A COMB .pep:*
4: /cgn2_6/ptodata/2/1aa/6B COMB .pep:*
5: /cgn2_6/ptodata/2/1aa/PTUS COMB .pep:*
6: /cgn2_6/ptodata/2/1aa/backfile1 .pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	260	93.9	260	4	US-09-536-647-2
2	45	16.2	277	4	US-09-536-647-3

ALIGNMENTS

RESULT 1
US-09-536-647-2
; Sequence 2, Application US/09536647
; Patent No. 6579708
; GENERAL INFORMATION:
; APPLICANT: Ho, Yen Sen
; APPLICANT: Johnson, Randall
; TITLE OF INVENTION: No. 6579708el Human Uridine Kinase
; FILE REFERENCE: GP50020
; CURRENT APPLICATION NUMBER: US/09/536,647
; CURRENT FILING DATE: 2000-03-27
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 260
; TYPE: PRT
; ORGANISM: Human
US-09-536-647-2

Query Match 93.9%; Score 260; DB 4; Length 260;

Best Local Similarity 100.0%; Pred. No. 4.6e-252;
Matches 260; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MASAGEDCESPAPEADRP...HPGLTRGKRSHLESSSRPH 277
DB	1	MASAGEDCESPAPEADRP...HPGLTRGKRSHLESSSRPH 277
QY	61	ILSODRFYKVLPAEQAKALKGQYNFDPDAFNDMHRTLKNIYEGKTVEPTYPVTH 120
DB	61	ILSODRFYKVLPAEQAKALKGQYNFDPDAFNDMHRTLKNIYEGKTVEPTYPVTH 120
QY	121	SRLEPTTVYPADVVLFEGLIVFYSGEIRDMFHLRLFVDTSDVRLSRVLRDVRGRDI 180
DB	121	SRLEPTTVYPADVVLFEGLIVFYSGEIRDMFHLRLFVDTSDVRLSRVLRDVRGRDI 180
QY	181	EQILTYITTFVKPAFEFCPLTKKYADVILIPRGVDMVAINILVQHIDILNDICKMR 240
DB	181	EQILTYITTFVKPAFEFCPLTKKYADVILIPRGVDMVAINILVQHIDILNDICKMR 240
QY	241	GGSGRSYKRTFSEPGDHPG 260
DB	241	GGSGRSYKRTFSEPGDHPG 260

RESULT 2
US-09-536-647-3
; Sequence 3, Application US/09536647
; Patent No. 6579708
; GENERAL INFORMATION:
; APPLICANT: Ho, Yen Sen
; APPLICANT: Johnson, Randall
; TITLE OF INVENTION: No. 6579708el Human Uridine Kinase
; FILE REFERENCE: GP50020
; CURRENT APPLICATION NUMBER: US/09/536,647
; CURRENT FILING DATE: 2000-03-27
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 277
; TYPE: PRT
; ORGANISM: Human
US-09-536-647-3

Query Match 16.2%; Score 45; DB 4; Length 277;
Best Local Similarity 100.0%; Pred. No. 5.4e-37;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	100	TLKNIYEGKTVEPTYPVTHSRLEPTTVYPADVVLFEGLIVFY 144
DB	100	TLKNIYEGKTVEPTYPVTHSRLEPTTVYPADVVLFEGLIVFY 144

Search completed: November 25, 2003, 08:11:26
Job time : 19 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2003, 08:09:10 ; Search time 31 seconds

(without alignments)
1648.090 Million cell updates/sec

Title: US-09-896-522-2

Sequence: 1 MASAGEDCEPAPADRP...HPGMLTSGKSHLESSSRPH 277

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 673684 seqs, 184443283 residues

Word size : 15

Total number of hits satisfying chosen parameters: 4

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

1: Published Applications AA.*
2: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09C_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	277	100.0	277	9	US-09-896-522-2
2	277	21.7	60	12	US-10-029-386-28765
3	25	9.0	337	10	US-09-925-300-1160
4	15	5.4	125	9	US-09-896-522-5

ALIGNMENTS

RESULT 1
US-09-896-522-2
; Sequence 2, Application US/09896522
; Patent No. US20020055161A1
; GENERAL INFORMATION:
; APPLICANT: Gluckemann, Maria A.

TITLE OF INVENTION: 57658, A NOVEL HUMAN URIDINE KINASE AND
FILE OF INVENTION: USES THEREOF
FILE REFERENCE: 381552001700
CURRENT APPLICATION NUMBER: US/09/896,522
CURRENT FILING DATE: 2001-06-28
PRIOR APPLICATION NUMBER: 60/216,503
PRIOR FILING DATE: 2000-06-30
NUMBER OF SEQ ID NOS: 6
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 277
TYPE: PRT
ORGANISM: Homo sapiens
US-09-896-522-2

Query Match 100.0%; Score 277; DB 9; Length 277;
Best Local Similarity 100.0%; Pred. No. 4e-264;
Matches 277; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASAGEDCEPAPADRP...HPGMLTSGKSHLESSSRPH 60
DB 1 MASAGEDCEPAPADRP...HPGMLTSGKSHLESSSRPH 60
QY 61 ILSODRFYKVLTAEGKAKALKGQYNFDPDAFNDIMRTLNIVEGKVEVPTVDFVTH 120
DB 61 ILSODRFYKVLTAEGKAKALKGQYNFDPDAFNDIMRTLNIVEGKVEVPTVDFVTH 120
QY 121 SRLPETTVVYPADVLFEGILVFYSQETIRDMHRLFVDTSDVRLSRRLVDVRGRDL 180
DB 121 SRLPETTVVYPADVLFEGILVFYSQETIRDMHRLFVDTSDVRLSRRLVDVRGRDL 180
QY 181 EQLLYQYTFYVPAFEEPLPTKYADVIIPRGVNMVAINIYOHIODILNGDICKMR 240
DB 181 EQLLYQYTFYVPAFEEPLPTKYADVIIPRGVNMVAINIYOHIODILNGDICKMR 240
QY 241 GGSNGRSYKRTFSEPDHFGMLTSGKSHLESSSRPH 277
DB 241 GGSNGRSYKRTFSEPDHFGMLTSGKSHLESSSRPH 277

RESULT 2
US-10-029-386-28765
; Sequence 28765, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR GI
; FILE REFERENCE: ABOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 28765
; LENGTH: 60
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR9.1
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.5
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.8
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.6
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.7
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3
; OTHER INFORMATION: EXPRESSED IN HEPA, SIGNAL = 3.7
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.3
; OTHER INFORMATION: SWISSPROT HIT: P52623, EVALU8 8.00e-26
US-10-029-386-28765

Query Match 21.7%; Score 60; DB 12; Length 60;

Best Local Similarity 100.0%; Pred. No. 3.1e-51;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 218 VAINLIYQHIQDILNGDICKWHRGSGNSYKRTFSEPGDHPGMLTSGKSHLESSSRPH 277
Db 1 VAINLIYQHIQDILNGDICKWHRGSGNSYKRTFSEPGDHPGMLTSGKSHLESSSRPH 60

RESULT 3

US-09-925-300-1160
; Sequence 1160, Application US/09925300
; Patent No. US2002015161A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1160
; LENGTH: 337
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (38)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (46)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (155)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (169)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-300-1160

Query Match 9.0%; Score 25; DB 10; Length 337;
Best Local Similarity 100.0%; Pred. No. 3.7e-16;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 189 TFWKPAFEFCLPTKXYADVVIIPRG 213
Db 263 TFWKPAFEFCLPTKXYADVVIIPRG 287

RESULT 4

US-09-896-522-5
; Sequence 5, Application US/09896522
; Patent No. US20020055161A1
; GENERAL INFORMATION:
; APPLICANT: Glucksmann, Maria A.
; TITLE OF INVENTION: 57658, A NOVEL HUMAN URIDINE KINASE AND
; FILE REFERENCE: 381552001700
; CURRENT APPLICATION NUMBER: US/09/896,522
; CURRENT FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: 60/216,503
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 125
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:

; OTHER INFORMATION: Consensus amino acid sequence
US-09-896-522-5

Query Match 5.4%; Score 15; DB 9; Length 125;
Best Local Similarity 100.0%; Pred. No. 1.1e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 246 RSYKRTFSEPGDHPG 260
Db 94 RSYKRTFSEPGDHPG 108

Search completed: November 25, 2003, 08:12:12
Job time : 32 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Comugen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2003, 08:07:30 ; Search time 19 Seconds
(without alignments)
1402.038 Million cell updates/sec

Title: US-09-896-522-2

Perfect score: 277

Sequence: 1 MASAGGDCESPAPAEADRP.....HPGMLTSGKRSHESSSRPH 277

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 283308 seqs, 96168682 residues

Word size : 15

Total number of hits satisfying chosen parameters: 0

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

.PIR 76:*
1: pirl:*
2: pirl:*
3: pirl:*
4: pirl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	Query	
No.	Score	Match Length DB ID Description

No matches found

Search completed: November 25, 2003, 08:10:52
Job time : 19 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2003, 08:06:20 ; Search time 11 Seconds

(without alignments)
1184.218 Million cell updates/sec

Title: US-09-896-522-2

Perfect score: 277

Sequence: 1 MASAGEDCEPAPADRP...HPGMLTSGKRSKSHLESSRPH 277

Scoring table: OLIGO

Gapop 60.0 , Gapept 60.0

Searched: 127863 seqs, 47026705 residues

Word size : 15

Total number of hits satisfying chosen parameters: 5

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

SUMMARIES

Result No. Score Query Length DB ID Description

1 277 100.0 277 1 UCK1_HUMAN

2 45 16.2 277 1 UCK1_MOUSE

3 25 9.0 261 1 UCK2_HUMAN

4 25 9.0 261 1 UCK2_MOUSE

5 15 5.4 260 1 UCK_DROME

ALIGNMENTS

RESULT 1

UCK1_HUMAN STANDARD; PRT; 277 AA.

AC O9HA47;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Uridine-cytidine kinase 1 (EC 2.7.1.48) (UCK 1) (Uridine

monophosphokinase 1) (Cytidine monophosphokinase 1).

GN UCK1.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A., AND CHARACTERIZATION.

RX MEDLINE=21203813; PubMed=11306702;

RA Van Rompay A.R., Norda A., Linden K., Johansson M., Karlsson A.;

RT "Phosphorylation of uridine and cytidine nucleoside analogs by two

human uridine-cytidine kinases.";

RL Mol. Pharmacol. 59:1181-1186(2001).

RN [2]

RP SEQUENCE FROM N.A.

RA Ho Y.S., Johnson R.K.;

RT "Human uridine kinase from prostate cancer cell line (LNCap).";

RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE FROM N.A.

RC TISSUE=Mammary gland;

RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,

RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,

RA Magatsuna M., Hosoki T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,

RA Takahashi M., Chida Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,

RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahara K., Masuno Y.,

RA Niinomiya K., Iwayanagi T.;

RL "NEO human cDNA sequencing project.";

RT Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.

RN [4]

RP SEQUENCE FROM N.A.

RA Xin Y.R., Yu L., Zhao S.Y.;

RT "Cloning of a new human cDNA similar to Mus musculus uridine kinase

mRNA.";

RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.

CC - FUNCTION: Phosphorylates uridine and cytidine to uridine

monophosphate and cytidine monophosphate. Does not phosphorylate

deoxyribonucleosides or purine ribonucleosides. Can use ATP or GTP

as a phosphate donor. Can also phosphorylate cytidine and uridine

nucleoside analogs such as 6-azauridine, 5-fluorouridine, 4-

thiouridine, 5-bromouridine, N(4)-acetylcytidine, N(4)-

benzoylcytidine, 5-fluorocytidine, 2-thiocytidine, 5-

methylecytidine, and N(4)-anticyclotidine.

CC - CATALYTIC ACTIVITY: ATP + uridine = ADP + UMP.

CC - 1- PATHWAY: Pyrimidine salvage pathway.

CC - 1- TISSUE SPECIFICITY: Ubiquitous.

CC - 1- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

or send an email to license@isb-sib.ch).

DR EMBL: AF237290; AAK28324.1; -

DR EMBL: AF254133; AAK49122.1; -

DR EMBL: AK022317; BAB14010.1; -

DR EMBL: AF125106; AAL75943.1; -

DR InterPro: IPR006083; PRK URK.

DR InterPro: IPR000764; Uridine_kin.

DR Pfam: PF00485; PRK. 1.

DR PRINTS: PR00986; URIDINKINASE.

DR TIGRPFAMs: TIGR00235; udk. 1.

KW transferase; Kinase; ATP-binding.

FT NP BIND 30 37

FT CONFILCT 8 17

FT DCESAPAEAD -> GABRAGAN (IN REF. 4).

FT CONFLICT 56 57

FT CONFLICT 247 247

FT SEQUENCE 277 AA; 31434 MW; AFD9ED92780CD502 CRC64;

Query Match 100.0%; Score 277; DB 1; Length 277;

Best local similarity 100.0%; Pred. No. 5.6e-276;

Matches 277; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASAGEDCEPAPADRP...HPGMLTSGKRSKSHLESSRPH 277

DB 1 MASAGEDCEPAPADRP...HPGMLTSGKRSKSHLESSRPH 277

QY 1 MASAGEDCEPAPADRP...HPGMLTSGKRSKSHLESSRPH 277

DB 1 MASAGEDCEPAPADRP...HPGMLTSGKRSKSHLESSRPH 277

QY 1 MASAGEDCEPAPADRP...HPGMLTSGKRSKSHLESSRPH 277

DB 1 MASAGEDCEPAPADRP...HPGMLTSGKRSKSHLESSRPH 277

QY 1 MASAGEDCEPAPADRP...HPGMLTSGKRSKSHLESSRPH 277

DB 1 MASAGEDCEPAPADRP...HPGMLTSGKRSKSHLESSRPH 277

QY 1 MASAGEDCEPAPADRP...HPGMLTSGKRSKSHLESSRPH 277

DB 1 MASAGEDCEPAPADRP...HPGMLTSGKRSKSHLESSRPH 277

QY 1 MASAGEDCEPAPADRP...HPGMLTSGKRSKSHLESSRPH 277

DB 1 MASAGEDCEPAPADRP...HPGMLTSGKRSKSHLESSRPH 277

QY 1 MASAGEDCEPAPADRP...HPGMLTSGKRSKSHLESSRPH 277

DB 1 MASAGEDCEPAPADRP...HPGMLTSGKRSKSHLESSRPH 277

QY 1 MASAGEDCEPAPADRP...HPGMLTSGKRSKSHLESSRPH 277

DB 1 MASAGEDCEPAPADRP...HPGMLTSGKRSKSHLESSRPH 277

QY 1 MASAGEDCEPAPADRP...HPGMLTSGKRSKSHLESSRPH 277

DB 1 MASAGEDCEPAPADRP...HPGMLTSGKRSKSHLESSRPH 277

QY 1 MASAGEDCEPAPADRP...HPGMLTSGKRSKSHLESSRPH 277

DB 1 MASAGEDCEPAPADRP...HPGMLTSGKRSKSHLESSRPH 277

```

Db      121 SRLETTVVYPADVLFEGILVFYSQELRDMFHLRLFPVDTSDVRLSRVLRDRRDL 180
Qy      181 EQLITQYTTFFVKPAPEEFCELPFKKYADVITPRGVDMVAIMLIYQHIDINDGICWHR 240
Db      181 EQLITQYTTFFVKPAPEEFCELPFKKYADVITPRGVDMVAIMLIYQHIDINDGICWHR 240
Qy      241 GGSNGRSYKRTFSEPDHPCGMLTSGKRSHESSSRPH 277
Db      241 GGSNGRSYKRTFSEPDHPCGMLTSGKRSHESSSRPH 277

RESULT 2
UCX1_MOUSE
ID      UCX1_MOUSE      STANDARD;      PRT;      277 AA.
AC      P52623;
DT      01-OCT-1996 (Rel. 34, Created)
DT      28-FEB-2003 (Rel. 41, Last sequence update)
DT      15-SEP-2003 (Rel. 42, Last annotation update)
DE      Uridine-cytidine kinase 1 (EC 2.7.1.48) (UCX 1) (Uridine
DE      monophosphokinase 1) (Cytidine monophosphokinase 1).
GN      UCX1 OR UMPK.
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxID=10090;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      TISSUE=Salivary gland;
RX      MEDLINE=22388257; PubMed=12477932;
RA      Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA      Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RA      Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA      Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA      Datcchenko L., Maruina K., Farmer A.A., Rubin G.M., Hong L.,
RA      Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA      Brownstein M.J., Uceda T.B., Toshiyuki S., Carrinci P., Prange C.,
RA      Raha S.S., Loguallano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA      Bosak S.A., McGowan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA      Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA      Villalón D.K., Muzny D.M., Sodexren E.J., Lu X., Gibbs R.A.,
RA      Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA      Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA      Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA      Butcherfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
RA      Schermer A., Schein J.E., Jones S.J.M., Marra M.A.;
RT      "Generation and initial analysis of more than 15,000 full-length
RT      human and mouse cDNA sequences.";
RL      Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN      [2]
RP      SEQUENCE OF 18-277 FROM N.A.
RC      TISSUE=Brain;
RX      MEDLINE=97108719; PubMed=8951040;
RA      Ropp P.A., Traut T.W.;
RT      "Cloning and expression of a cDNA encoding uridine kinase from mouse
RT      brain.";
RL      Arch. Biochem. Biophys. 336:105-112(1996).
CC      -1- FUNCTION: Phosphorylates uridine and cytidine to uridine
CC      monophosphate and cytidine monophosphate. Does not phosphorylate
CC      deoxyribonucleosides or purine ribonucleosides. Can use ATP or GTP
CC      as a phosphate donor. Can also phosphorylate cytidine and uridine
CC      nucleoside analogs such as 6-azauridine, 5-fluorouridine, 4-
CC      thioridine, 5-bromouridine, N(4)-acetylcytidine, N(4)-
CC      benzoylcytidine, 5-fluorocytidine, 2-thiocytidine, 5-
CC      methylcytidine, and N(4)-anisoylcytidine (By similarity).
CC      -1- CATALYTIC ACTIVITY: ATP + uridine = ADP + UMP.
CC      -1- CATALYTIC ACTIVITY: ATP + cytidine = ADP + CMP.
CC      -1- PATHWAY: Pyrimidine salvage pathway.
CC      -1- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its

```

```

CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; BC025146; AAH25146.1; -.
DR      EMBL; L31783; AAB50568.1; -.
DR      MGI; MGI:98804; Umpk.
DR      InterPro; IPR006083; PRK_UK.
DR      InterPro; IPR00764; Uridine_kin.
DR      Pfam; PF00485; PRK_1.
DR      PRINTS; PR00988; URIDINKINASE.
DR      TIGRFAMs; TIGR00235; udk; 1.
KW      Transferase; Kinase; ATP-binding.
FT      NP_BIND 30
SQ      SEQUENCE 277 AA; 31068 MW; 3EBB3C4187FAEB4A CRC64;

Query Match      16.2%; Score 45; DB 1; Length 277;
Best Local Similarity 100.0%; Pred. No. 2,9e-38;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      100 TLKNIVEKTEVEPTVDFTVTHSRLEPTTVVYPADVLFEGILVFY 144
Db      100 TLKNIVEKTEVEPTVDFTVTHSRLEPTTVVYPADVLFEGILVFY 144

RESULT 3
UCX2_HUMAN
ID      UCX2_HUMAN      STANDARD;      PRT;      261 AA.
AC      Q9B2X2; Q96RG5;
DT      28-FEB-2003 (Rel. 41, Created)
DT      28-FEB-2003 (Rel. 41, Last sequence update)
DT      28-FEB-2003 (Rel. 41, Last annotation update)
DE      Uridine-cytidine kinase 2 (EC 2.7.1.48) (UCX 2) (Uridine
DE      monophosphokinase 2) (Cytidine monophosphokinase 2).
GN      UCX2 OR UMPK.
OS      Homo sapiens (human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX      NCBI_TaxID=9606;
RN      [1]
RP      SEQUENCE FROM N.A.. AND CHARACTERIZATION.
RX      MEDLINE=21203813; PubMed=11306702;
RA      Van Rompay A.R., Norda A., Linden K., Johansson M., Karlsson A.;
RT      "Phosphorylation of uridine and cytidine nucleoside analogs by two
RT      human uridine-cytidine kinases.";
RL      Mol. Pharmacol. 59:1181-1186(2001).
RN      [2]
RP      SEQUENCE OF 15-261 FROM N.A.
RC      TISSUE=Fibrosarcoma;
RX      MEDLINE=21385121; PubMed=11494055;
RA      Koizumi K., Shimamoto Y., Azuma A., Mataya Y., Matsuda A., Sasaki T.,
RA      Fukushima M.;
RT      "Cloning and expression of uridine/cytidine kinase cDNA from human
RT      fibrosarcoma cells.";
RL      Int. J. Mol. Med. 8:273-278(2001).
CC      -1- FUNCTION: Phosphorylates uridine and cytidine to uridine
CC      monophosphate and cytidine monophosphate. Does not phosphorylate
CC      deoxyribonucleosides or purine ribonucleosides. Can use ATP or GTP
CC      as a phosphate donor. Can also phosphorylate cytidine and uridine
CC      nucleoside analogs such as 6-azauridine, 5-fluorouridine, 4-
CC      thioridine, 5-bromouridine, N(4)-acetylcytidine, N(4)-
CC      benzoylcytidine, 5-fluorocytidine, 2-thiocytidine, 5-
CC      methylcytidine, and N(4)-anisoylcytidine.
CC      -1- CATALYTIC ACTIVITY: ATP + uridine = ADP + UMP.
CC      -1- CATALYTIC ACTIVITY: ATP + cytidine = ADP + CMP.
CC      -1- PATHWAY: Pyrimidine salvage pathway.
CC      -1- TISSUE SPECIFICITY: Expressed in placenta.
CC      -1- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its

```


CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 CC EMBL; AF236637; AKL14053.1; -
 CC EMBL; AB062451; BAB56162.1; -
 CC Genew; HGNC:12562; UMPK.
 CC InterPro; IPR006082; PRK.
 CC InterPro; IPR006083; PRK_URK.
 CC InterPro; IPR000764; Uridine_kin.
 CC Pfam; PF00485; PRK.1.
 CC PRINTS; PR00478; PHRIBLKINASE.
 CC PRINTS; PR00988; URIDINKINASE.
 CC TIGRFAme; TIGR00235; udk.1.
 CC Transferase; Kinase; ATP-binding.
 CC NP_BIND 27 34 ATP (POTENTIAL).
 CC SEQUENCE 261 AA; 29299 MW; 71791346F091EBFD CRC64;
 SQ

Query Match 9.0%; Score 25; DB 1; Length 261;
 Best Local Similarity 100.0%; Pred. No. 8.5e-18;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 189 TFWKPAFEFECUPTKKYADVILPRG 213
 |||||
 DB 187 TFWKPAFEFECUPTKKYADVILPRG 211

RESULT 4
 UCK2 MOUSE STANDARD; PRT; 261 AA.
 AC 099PM9;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Uridine-cytidine kinase 2 (EC 2.7.1.48) (UCK 2) (Uridine
 DE monophosphokinase 2) (Cytidine monophosphokinase 2).
 GN UCK2 OR UMPK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_Taxid=10090;
 [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21203813; PubMed=11306702;
 RA Van Rompay A.R., Norda A., Lindén K., Johansson M., Karlsson A.;
 RT "Phosphorylation of uridine and cytidine nucleoside analogs by two
 RT human uridine-cytidine kinases.";
 RT Mol. Pharmacol. 59:1181-1186(2001).
 RL

CC -1- FUNCTION: Phosphorylates uridine and cytidine to uridine
 CC monophosphate and cytidine monophosphate. Does not phosphorylate
 CC deoxyribonucleosides or purine ribonucleosides. Can use ATP or GTP
 CC as a phosphate donor. Can also phosphorylate cytidine and uridine
 CC nucleoside analogs such as 6-azauridine, 5-fluorouridine, 4-
 CC thioridine, 5-bromouridine, N(4)-acetylcytidine, N(4)-
 CC benzoylcytidine, 5-fluorocytidine, 2-thiocytidine, 5-
 CC methylcytidine, and N(4)-aminoacylcytidine (By similarity).
 CC -1- CATALYTIC ACTIVITY: ATP + uridine = ADP + UMP.
 CC -1- PATHWAY: Pyrimidine salvage pathway.
 CC -1- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 CC EMBL; AF236636; AKL14052.1; -
 CC MGD; MGI:1931744; Uck2.
 CC InterPro; IPR006082; PRK.

DR InterPro; IPR006083; PRK_URK.
 DR InterPro; IPR000764; Uridine_kin.
 DR Pfam; PF00485; PRK.1.
 DR PRINTS; PR00478; PHRIBLKINASE.
 DR PRINTS; PR00988; URIDINKINASE.
 DR TIGRFAme; TIGR00235; udk.1.
 DR Transferase; Kinase; ATP-binding.
 DR NP_BIND 27 34 ATP (POTENTIAL).
 DR SEQUENCE 261 AA; 29404 MW; 780AA3CF5CA8153 CRC64;
 SQ

Query Match 9.0%; Score 25; DB 1; Length 261;
 Best Local Similarity 100.0%; Pred. No. 8.5e-18;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 189 TFWKPAFEFECUPTKKYADVILPRG 213
 |||||
 DB 187 TFWKPAFEFECUPTKKYADVILPRG 211

RESULT 5
 UCK DROME STANDARD; PRT; 260 AA.
 ID UCK_DROME
 AC 09VC99;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Probable uridine-cytidine kinase (EC 2.7.1.48) (UCK) (Uridine
 DE monophosphokinase) (Cytidine monophosphokinase).
 GN CG6364.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_Taxid=7227;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkely;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
 RA Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Aguirre A., An H.-J., Andrews-Pfannkoch C., Balowin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Bertman B.P., Bhargava D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brocktein P., Brodtier P.,
 RA Burks K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Folsch C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glaeser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Kethum K.A.,
 RA Kimball B.E., Kodira C.D., Kraft C., Kravitz S., Kuip D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusslein-Dhar B., Pacle J.M.,
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puti V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skurski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveril J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,

RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*,";
 RL Science 287:2185-2195(2000).
 CC -1- CATALYTIC ACTIVITY: ATP + uridine = ADP + UMP.
 CC -1- CATALYTIC ACTIVITY: ATP + cytidine = ADP + CMP.
 CC -1- PATHWAY: Pyrimidine salvage pathway.
 CC -1- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AE003747; AAF56274.1; -
 DR P1yBase; FBgn0039179; CG6364.
 DR InterPro; IPR006083; PRK_UK.
 DR InterPro; IPR000764; Uridine_kin.
 DR Pfam; PF00485; PRK_1.
 DR PRINTS; PR00988; URIDINKINASE.
 KW Hypothetical protein; Transferase; Kinase; ATP-binding.
 FT NP BIND 34 41 ATP (POTENTIAL).
 SO SEQUENCE 260 AA; 29213 MW; 2EE57423704925E6 CRC64;

Query Match 5.4%; Score 15; DB 1; Length 260;
 Best Local Similarity 100.0%; Pred. No. 1.5e-07;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 130 YPADVLFEGILVY 144
 |||||
 Db 135 YPADVLFEGILVY 149

Search completed: November 25, 2003, 08:09:29
 Job time : 11 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2003, 08:07:00 ; Search time 35 Seconds
(without alignments)
2042.302 Million cell updates/sec

Title: US-09-896-522-2
Perfect score: 277
Sequence: 1 MASAGEDSPAPADRP.....HPGMLTSGKSHLESSSRPH 277

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 830525 seqs, 258052604 residues

Word size : 15

Total number of hits satisfying chosen parameters: 6

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :
1: SP archaea:*
2: SP bacteria:*
3: SP fungi:*
4: SP human:*
5: SP invertebrate:*
6: SP mammal:*
7: SP mhc:*
8: SP organelle:*
9: SP phage:*
10: SP plant:*
11: SP rodent:*
12: SP virus:*
13: SP vertebrate:*
14: SP unclassified:*
15: SP viirus:*
16: SP bacteriap:*
17: SP archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	169	61.0	201	4 Q96B0	Q96B0 homo sapien
2	25	9.0	105	11 Q9QY8	Q9QY8 rattus norv
3	25	9.0	111	4 Q9BU42	Q9BU42 homo sapien
4	25	9.0	261	11 Q8C476	Q8C476 mus musculi
5	15	5.4	111	4 Q92528	Q92528 homo sapien
6	15	5.4	260	5 Q8MRJ1	Q8MRJ1 drosophila

ALIGNMENTS

RESULT 1
Q96B0 PRELIMINARY; PRT; 201 AA.
AC Q96B0; DT 01-DEC-2001 (TReMBLrel. 19, Created)

DT 01-DEC-2001 (TReMBLrel. 19, last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, last annotation update)
DE Similar to uridine-cytidine kinase 1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Colon;
RA Strausberg R.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC015547; AAH15547.1; -
DR InterPro; IPR006083; PRK URK.
DR InterPro; IPR000764; Uridine_kin.
DR Pfam; PF00465; PRK.1.
DR PRINTS; PR00988; URIDINKINASE.
KW Kinase.
SQ SEQUENCE 201 AA; 22790 MW; 0E5F2F00F7B5363 CRC64;

Query Match 61.0%; Score 169; DB 4; Length 201;
Best Local Similarity 100.0%; Pred. No. 2.1e-170; Indels 0; Gaps 0;
Matches 169; Conservative 0; Mismatches 0;

QY 1 MASAGEDSPAPADRP...PFLIGVSGTASGKSTVCEKIMELLGONEVEQORQKV 60
DB 1 MASAGEDSPAPADRP...PFLIGVSGTASGKSTVCEKIMELLGONEVEQORQKV 60
QY 61 IISQDRFYKVLTAEOAKALKQYNFDPDAPNDIMHRTLNIVGKTVETVYDFVTH 120
DB 61 IISQDRFYKVLTAEOAKALKQYNFDPDAPNDIMHRTLNIVGKTVETVYDFVTH 120
QY 121 SRLETTVVYPADVVFEGILVYSGEIRDMHRLRFVYDSDVRSRR 169
DB 121 SRLETTVVYPADVVFEGILVYSGEIRDMHRLRFVYDSDVRSRR 169

RESULT 2
Q9QY8 PRELIMINARY; PRT; 105 AA.
AC Q9QY8;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, last sequence update)
DT 01-OCT-2002 (TReMBLrel. 22, last annotation update)
DE Uridine kinase (Fragment).
GN UK.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SD; TISSUE=spinal cord;
RX MEDLINE=20050059; PubMed=10581173;
RA Yuh I., Yao T., Watanabe S., Okajima S., Hirasawa Y., Fushiki S.;
RT "Up-regulated uridine kinase gene identified by RLCs in the ventral
horn after crush injury to rat sciatic nerves."
RL Biochem. Biophys. Res. Commun. 266:104-109(1999).
DR EMBL; AB030700; BA83085.1; -
DR InterPro; IPR006083; PRK URK.
DR InterPro; IPR000764; Uridine_kin.
DR Pfam; PF00465; PRK.1.
DR PRINTS; PR00988; URIDINKINASE.
KW Kinase.
FT NON_TER 1
FT NON_TER 105
SQ SEQUENCE 105 AA; 11967 MW; 9DB43C15E07EB029 CRC64;

Query Match 9.0%; Score 25; DB 11; Length 105;
Best Local Similarity 100.0%; Pred. No. 4.6e-18;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 169 TFVKPAFEFFCLPTKKYADVILIRG 213

Db 36 TFVKPAFEFCLPTKXADVIIPRG 60

RESULT 3

Q9BU42 PRELIMINARY; PRT; 111 AA.
 AC Q9BU42; 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE Similar to uridine monophosphate kinase.
 OS Homo sapiens (Human)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Lung;
 RA Strauberg R.;
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC002906; AA02906.1; -
 DR InterPro; IPR006083; PRK URK.
 DR InterPro; IPR00764; Uridine_kin.
 DR Pfam; PF00485; PRK.1.
 DR PRINTS; PR00988; URIDINKINASE.
 KW Kinase.
 SQ SEQUENCE 111 AA; 12587 MW; E6688B186FA32A9 CRC64;

Query Match 9.0%; Score 25; DB 4; Length 111;
 Best Local Similarity 100.0%; Pred. No. 4.8e-18;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 189 TFVKPAFEFCLPTKXADVIIPRG 213
 Db 37 TFVKPAFEFCLPTKXADVIIPRG 61

RESULT 4

Q8C476 PRELIMINARY; PRT; 261 AA.
 AC Q8C476; 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Uridine-cytidine kinase 2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J;
 RX MEDLINE=22354683; PubMed=12466851;
 RA The FANTOM Consortium,
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.";
 RL Nature 420:563-573 (2002).
 DR EMBL; AK082837; BAC38646.1; -
 SQ SEQUENCE 261 AA; 29438 MW; 780AA3C298AA8153 CRC64;

Qy 189 TFVKPAFEFCLPTKXADVIIPRG 213
 Db 187 TFVKPAFEFCLPTKXADVIIPRG 211

RESULT 5

Q92528

ID Q92528 PRELIMINARY; PRT; 111 AA.

AC Q92528; 01-FEB-1997 (TrEMBLrel. 02, Created)
 DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE 5'-terminal region of UMK.
 OS Homo sapiens (Human)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Testis;
 RX MEDLINE=96411689; PubMed=8912458;
 RA Oaki K., Kuraki T., Hayashi S., Nakamura Y.;
 RT "Isolation of three testis-specific genes (TSA303, TSA806, TSA903) by
 RT a differential mRNA display method.";
 RL Genomics 36:316-319 (1996).
 DR EMBL; D78335; BA011349.1; -
 DR InterPro; IPR006083; PRK URK.
 DR InterPro; IPR00764; Uridine_kin.
 DR Pfam; PF00485; PRK.1.
 DR PRINTS; PR00988; URIDINKINASE.
 SQ SEQUENCE 111 AA; 12617 MW; 81F6E4019CF428AB CRC64;

Query Match 5.4%; Score 15; DB 4; Length 111;
 Best Local Similarity 100.0%; Pred. No. 1.9e-07;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 189 TFVKPAFEFCLPTK 203
 Db 37 TFVKPAFEFCLPTK 51

RESULT 6

Q8MRJ1 PRELIMINARY; PRT; 260 AA.
 AC Q8MRJ1; 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE LD13909P.
 GN CG6364.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkley;
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
 RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,
 RA Miranda A., Mungall C.J., Nuno J., Pacleb J., Parasas V., Park S.,
 RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
 RA Ceiniker S.;
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY119583; AAM50237.1; -
 DR FlyBase; FBgn0039179; CG6364.
 DR InterPro; IPR006083; PRK URK.
 DR InterPro; IPR00764; Uridine_kin.
 DR Pfam; PF00485; PRK.1.
 DR PRINTS; PR00988; URIDINKINASE.
 SQ SEQUENCE 260 AA; 29223 MW; 03CAA6DCA04A16E5 CRC64;

Query Match 5.4%; Score 15; DB 5; Length 260;
 Best Local Similarity 100.0%; Pred. No. 4.1e-07;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 130 YPADVLFEGILVY 144
 Db 135 YPADVLFEGILVY 149

Wed Nov 26 18:16:18 2003

us-09-896-522-2.011.rspc

Page 3

Search completed: November 25, 2003, 08:10:18
Job time : 35 secs

This Page Blank (uspto)

Db 241 GlycylseranGlyArgSerTyrIyArgThrPheSerGluProGlyAspHisProGly 260

RESULT 2

US-09-536-647-3
; Sequence 3, Application US/09536647
; Patent No. 6579708
; GENERAL INFORMATION:
; APPLICANT: Ho, Yen Sen
; APPLICANT: Johnson, Randall
; TITLE OF INVENTION: No. 6579708e1 Human Uridine Kinase
; FILE REFERENCE: GP50020
; CURRENT APPLICATION NUMBER: US/09/536,647
; CURRENT FILING DATE: 2000-03-27
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 277
; TYPE: PRT
; ORGANISM: Human
US-09-536-647-3

Alignment Scores:

Pred. No.:	1.38e-34	Length:	277
Score:	45.00	Matches:	45
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	8.40%	Indels:	0
DB:	4	Gaps:	0

US-09-896-522-1 (1-1624) x US-09-536-647-3 (1-277)

QY	391	ACTCTGAGAAACATCGTGGAGGGCAAAACGGTGGAGTGCCGACCTATGATTTGTGACA	450
Db	100	ThrLeuIysAsnIleValGluGlyThrValGluValProThrTyrAspPheValThr	119
QY	451	CACTCAAGGTACCAAGACAGACGATGCTACCCCTGCGACGTGTTCTGTTGAGGGC	510
Db	120	HisSerArgLeuProGluThrThrValValTyrProIleAspValValLeuPheGluGly	139
QY	511	ATCTTGTTGTTCTAC	525
Db	140	IleLeuValPheTyr	144

Search completed: November 25, 2003, 08:01:00
Job time : 19.5085 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OW nucleic - protein search, using frame_plus_n2p model

Run on: November 25, 2003, 07:53:25 ; Search time 56.8202 Seconds
(without alignments)
10543.293 Million cell updates/sec

Title: US-09-896-522-1
Perfect score: 536
Sequence: 1 gtgggggtgcctccgcgaccc.....ccaggtctgtgcggggccag 1624

Scoring table:
OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 673684 seqs, 184443283 residues

Word size: 15

Total number of hits satisfying chosen parameters: 5

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODE=frame-n2p.model -DEV=xlh
-Q=/cgn2_1/USPTO.epool/US09896522/runat_21112003_184147_3282/app_query.fasta.1_2830
-DB=Published Applications_AA -QFMT=fastan -SUFFIX=n2p_ol1.rapb -MINMATCH=0.1
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR SCORE=quality -THR MIN=15
-ALIGN=15 -MODE=LOCAL -OUTFMT=ptco -NORM=ext -HEADSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US09896522 @CGN_1_1_129 @runat_21112003_184147_3282
-NCPU=6 -ICPU=3 -NO_MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSEBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=60 -YGAEXT=60 -DELOP=6 -DELEXT=7

Database :

Published Applications_AA: *
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep: *
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep: *
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep: *
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep: *
5: /cgn2_6/ptodata/1/pubpaa/PCTUS_NEW_PUB.pep: *
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep: *
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep: *
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep: *
9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep: *
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep: *
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep: *
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep: *
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep: *
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep: *
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep: *
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep: *
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep: *
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result Query #
No. Score Match length DB ID Description

1 277 51.7 277 9 US-09-896-522-2 Sequence 2, Appl1
2 60 11.2 60 12 US-10-029-386-28765 Sequence 28765, A
3 25 4.7 337 10 US-09-925-300-1160 Sequence 1160, Ap
4 19 3.5 73 10 US-09-764-877-1422 Sequence 1422, Ap
5 15 2.8 125 9 US-09-896-522-5 Sequence 5, Appl1

ALIGNMENTS

RESULT 1
US-09-896-522-2
; Sequence 2, Application US/09896522
; Patent No. US2002005161A1
; GENERAL INFORMATION:
; APPLICANT: Glucksmann, Maria A.
; TITLE OF INVENTION: 57658, A NOVEL HUMAN URIDINE KINASE AND
; FILE REFERENCE: 381552001700
; CURRENT APPLICATION NUMBER: US/09/896,522
; CURRENT FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: 60/216,503
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 277
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-896-522-2

Alignment Scores:

Score: 1 436-252 Length: 277
Percent Similarity: 277.00 Matches: 277
Best Local Similarity: 100.00% Conservative: 0
Query Match: 51.68% Mismatches: 0
DB: 9 Indels: 0
Gaps: 0

US-09-896-522-1 (1-1624) x US-09-896-522-2 (1-277)

QY 94 ATGGCTTCGGCGGAGGCGAAGACTGCGAGAGCCCGCGGAGCCGACCGTCGAC 153
DB 1 MetalaSerAlaGlyGlyGlnuSpCysglnuSerProAlaProGluAlaAspArgProHis 20
QY 154 CAGCGCCCTTCTGATAGGGGTGAGCGGCGCACTGCCAGCGGGAAGTCCAGCTGT 213
DB 21 GlnArgProPheLeuIleGlyValSerGlyThrAlaSerGlyLysSerThrValCys 40
QY 214 GAGAAGATCATGAGTTGCTGGGAGACGAGCGGAGAAAGCGGAGCGGAGGAGGTC 273
DB 41 GlnuPrlleuMetGlnuLeuLeuGlnuInuGlnuValGlnuGlnuArgGlnuValVal 60
QY 274 ATCTGAGGCGAGAGAGGTTCTCAAGGTCCTGACGCGAGACGAGAGGCGAAGGCTTG 333
DB 61 IleuSerGlnuAspArgPheThrLysValLeuThrAlaGlnuInuValAlaLysAlaLeu 80
QY 334 AAGGACGATACATTTTGACCATCCAGATGCTTGAATGATTTGATGACAGACT 393
DB 81 LysGlyGlnuTyAsnPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThr 100
QY 394 CTGAAGACATGCTGAGAGGCGAAACGAGTGAAGCGGAGCGGAGCTTATGATTTGTGACAC 453
DB 101 LeuLysAsnIleValGlnuGlyLysThrValGlnuAlaProThrTyAspPheValThrHis 120
QY 454 TCAGGTTACAGAGACACGAGTGTCTACCTGCGGAGCGTGTCTGTTGAGGCGATC 513
DB 121 SerArgLeuProGluThrThrValValTyProAlaAspValValLeuPheGlnuGlyLe 140
QY 514 TTGGTGTCTACAGCGGAGAGATCCGGACATGTTCCACTCGGCGCTCTTGCGACAC 573
DB 141 LeuValIlePheTySerGlnuGlnuIleArgAspMetPheHisLeuArgLeuPheValAspThr 160
QY 574 GACTCGAGAGTCAGGCTGTCTCGAAGAGTTCTCCGAGAGCGTCCGAGGAGGAGGACCTG 633

Db 161 AsperasperaValaGlueSerArgValaLeuArgAspValaGagGlyArgAspLeu 180
QY 634 GAGCAGATTCTGACGAGTACACACCTTCGTAAGCCGCTTCGAGAGTTCTGCTG 693
Db 181 GluGlnIleLeuThrGlnIleThrThrPheValysProAlaPheGluGluPheCysLeu 200
QY 694 CCGACAAAGAGTATGCCGATGTATCATCCACAGAGAGTGGACATATGTTGCCATC 753
Db 201 ProThrIysIysIyrAlaAspValaIleIleProArgGlyValaAspAsnMetValaIle 220
QY 754 AACCTGATCGTGCAGCATCCAGACATTTGATGATGTGACATCTGCAATGGACCGA 813
Db 221 AsnLeuIleValaGlnHisIleGlnAspIleLeuAsnGlyAspIleCysIysTTPHisArg 240
QY 814 GAGAGGTCCAAATGGCGGAGCTACAGCGACCTTTCTGAGCCAGGAGACACCTCTGG 873
Db 241 GlyIysIerAsnGlyArgSerTyrIysArgThrPheSerGluProGlyAspHisProGly 260
QY 874 ATGCTGACCTTGCGCAACGCTCACATTTGAGAGTCCAGCAGACACCCAC 924
Db 261 MetLeuThrSerGlyIysArgSerHisLeuGlnSerSerArgProHis 277

RESULT 2

US-10-029-386-28765
; Sequence 28765, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR C
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 28765
; LENGTH: 60
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR9.1
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.5
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.8
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.6
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.7
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.5
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.7
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.3
; OTHER INFORMATION: SWISSPROT HIT: P52623, EVALUE 8.00e-26
US-10-029-386-28765

Alignment Scores:

Pred. No.: 1.85e-47 Length: 60
Score: 60.00 Matches: 60
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 11.19% Indels: 0
DB: 12 Gaps: 0

US-09-896-522-1 (1-1624) x US-10-029-386-28765 (1-60)

QY 745 GTTCCATCAACCTGATCGTACACATCCAGACATTTGATGATGATGATGCAAA 804
Db 1 ValAlaIleAsnLeuIleValaGlnHisIleGlnAspIleLeuAsnGlyAspIleCysLys 20
QY 805 TGGCAGCAGAGAGGCTCCATATGGCGGAGACTACAGCGGACCTTTTCTGACCGAGGGGAC 864
Db 21 TrpHisArgGlyIysIerAsnGlyArgSerTyrIysArgThrPheSerGluProGlyAsp 40

QY 865 CACCTGGAGTGTGCTCTGTGGCAACGCTGACATTTGAGTCCAGACAGACCCAC 924
Db 41 HisProGlyMetLeuThrSerGlyIysArgSerHisLeuGlnSerSerArgProHis 60

RESULT 3

US-09-925-300-1160
; Sequence 1160, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben,
; APPLICANT: PA101
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1160
; LENGTH: 337
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (38)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (46)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (155)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (169)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-300-1160

Alignment Scores:

Pred. No.: 1.76e-14 Length: 337
Score: 25.00 Matches: 25
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 4.66% Indels: 0
DB: 10 Gaps: 0

US-09-896-522-1 (1-1624) x US-09-925-300-1160 (1-337)

QY 658 ACCTTGTTGTAAGCGCGCTTCGAGAGTCTGTGCTGCCGACAAAGATATGCCATGTG 717
Db 263 ThrPheValysProAlaPheGluGluPheCysLeuProThrIysIyrAlaAspVal 282

QY 718 ATCATCCACAGAGA 732
Db 283 IleIleProArgGly 287

RESULT 4

US-09-764-877-1422
; Sequence 1422, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005
; CURRENT APPLICATION NUMBER: US/09/764,877
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1422
; LENGTH: 73

; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (30)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (72)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-877-1422

Alignment Scores:

Pred. No.:	9.87e-09	Length:	73
Score:	19.00	Matches:	19
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.54%	Indels:	0
DB:	10	Gaps:	0

US-09-896-522-1 (1-1624) x US-09-764-877-1422 (1-73)

QY 1521 CCACTGAGGAGATCTTTCCCTTGCTTGCTTAAGGCAAGAGGAGCGG 1577

Db 11 ProleuArgAspValPheProCysLeuCyLeuLysGlyArgGluAaArg 29

RESULT 5

US-09-896-522-5

; Sequence 5, Application US/09896522
; Patent No. US20020055161A1

; GENERAL INFORMATION:

; APPLICANT: Glucksmann, Maria A.

; TITLE OF INVENTION: 57658, A NOVEL HUMAN URIDINE KINASE AND

; TITLE OF INVENTION: USRS, THEREOF

; FILE REFERENCE: 381552001700

; CURRENT APPLICATION NUMBER: US/09/896,522

; PRIOR FILING DATE: 2001-06-28

; PRIOR APPLICATION NUMBER: 60/216,503

; NUMBER OF SEQ ID NOS: 6

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 5

; LENGTH: 125

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Consensus amino acid sequence

US-09-896-522-5

Alignment Scores:

Pred. No.:	5.57e-05	Length:	125
Score:	15.00	Matches:	15
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.80%	Indels:	0
DB:	9	Gaps:	0

US-09-896-522-1 (1-1624) x US-09-896-522-5 (1-125)

QY 829 CCGAGCTACAGCGACCTTTCTGAGCCAGGGGACACCCCTGG 873

Db 94 ArgSerTyrIysArgThrPheSerGluProGlyAspHisProGly 108

Search completed: November 25, 2003, 08:04:07
Job time : 58.8202 secs

This Page Blank (uspto)

FEATURES

source

Location/Qualifiers

1. 1118
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5532885"
 /issue_type="leiomysarcoma"
 /lab_host="DH10B (phage-resistance)"
 /clone_id="NH_MGC_71"
 /note="Organ: uterus; Vector: pCMVSPORT6, Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
 Average insert size 2.1 kb."

BASE COUNT 234 a 311 c 361 g 209 t 3 others

ORIGIN

Query Match 54.9%; Score 890.8; DB 12; Length 1118;
 Best Local Similarity 90.9%; Pred. No. 3.6e-179;
 Matches 1005; Conservative 0; Mismatches 89; Indels 12; Gaps 5;

49 GAAGGGCGGGCGGGGACCCGATGCGGGGACCGGAGCCGAGATGCTTCGGCGGA 108
 2 GGAGCGGTGGCGGGGACCCGATGCGGGGACCGGAGCCGAGATGCTTCGGCGGA 61
 109 GCGGAAGACTGCGAGAGCCCGCGCGGAGCCGAGCCGAGATGCTTCGGCGGA 168
 62 GCGGAAGACTGCGAGAGCCCGCGCGGAGCCGAGCCGAGATGCTTCGGCGGA 121
 169 ATAGGGGTGAGCGCGGCACTGCGAGCGGAGATGCACTGTGTGAGAAAGATCATGAG 228
 122 ATAGGGGTGAGCGCGGCACTGCGAGCGGAGATGCACTGTGTGAGAAAGATCATGAG 181
 229 TTGCTGGAGACAGACGAGTGGAGACAGCGGAGCGGAGTGTGATCTGAGCAGAGAC 288
 182 TTGCTGGAGACAGACGAGTGGAGACAGCGGAGCGGAGTGTGATCTGAGCAGAGAC 241
 289 AGGTTCTACAGGTCCTGACGCGGAGACAGAGGCGGAGGCTTGAAGAGACAGTAAT 348
 242 AGGTTCTACAGGTCCTGACGCGGAGACAGAGGCGGAGGCTTGAAGAGACAGTAAT 301
 349 TTGACCATCCAGATGCTTGTATGATGATTTGATGCAAGAGACTTGAAGACATGCTG 408
 302 TTGACCATCCAGATGCTTGTATGATGATTTGATGCAAGAGACTTGAAGACATGCTG 361
 409 GAGGGCAAAACGAGGAGGCGGAGCCTATGATTTTGTGACACCTGAAGTTACAGAG 468
 362 GAGGGCAAAACGAGGAGGCGGAGCCTATGATTTTGTGACACCTGAAGTTACAGAG 421
 469 ACCAGGTGTCCTGACGCGGAGGCGGAGCCTATGATTTTGTGACACCTGAAGTTACAG 528
 422 ACCAGGTGTCCTGACGCGGAGGCGGAGCCTATGATTTTGTGACACCTGAAGTTACAG 481
 529 CAGAGATCCGAGACATGCTTCACCTGCGCTCTTCTGACACCGACTCCAGAGTCAAG 588
 482 CAGAGATCCGAGACATGCTTCACCTGCGCTCTTCTGACACCGACTCCAGAGTCAAG 541
 589 CTGCTGGAAGAGTTCCTCGGAGAGTGCCTCGAGGAGGAGCCTGAGACATTCAGAG 648
 542 CTGCTGGAAGAGTTCCTCGGAGAGTGCCTCGAGGAGGAGCCTGAGACATTCAGAG 601
 649 CAGTACACCACTTCGTAAGCCGCGCTTCGAGAGGTTCTGCGGACCAAGAAATAT 708
 602 CAGTACACCACTTCGTAAGCCGCGCTTCGAGAGGTTCTGCGGACCAAGAAATAT 661
 709 GCGCATGTATCATCCACGAGAGTGAACAATATGTTGTCATCAACTGATGTGAG 768
 662 GCGCATGTATCATCCACGAGAGTGAACAATATGTTGTCATCAACTGATGTGAG 721
 769 CAGTACAGAGCATTCGAATGTGACATTCGAATGTGACCAAGAGAGGTCGAATGG 828
 722 CAGTACAGAGCATTCGAATGTGACATTCGAATGTGACCAAGAGAGGTCGAATGG 781
 829 CGAGACTACAGCGGAGCTTTCTGAGCAGGAGGAGCAGCTTG-GATGCTGACCTTGG 887

Db 782 CGAGGTACAAGGGAAGCTTTCTGAGCAGGAGGAGCAGCCTGAGTGTACCTCTGG 841
 Qy 888 CAACAGGTACATTTGAGATCCAGAGCAGA-CCCACTGAGGGGCTGCCAGCCTCAG 946
 Db 842 CAACAGGTACATTTGAGATCCAGAGCAGAGACCCCACTGAGGGGCTGCCAGCCTCAG 901
 Qy 947 GCAAGTCTCCCGCGG---CATGTGTTCAGGAGCTAGAGCTGGGAGCGCCACCCAG 1003
 Db 902 GCGAGTCTTGAGTCCCGGAGTGGGGGTTTCCGAACTGAACCTGGGAGAGCCACCA 961
 Qy 1004 ACCCACTGCTCTCTCTCGGAGCAGCC---AGGAGGTGTGACAGAGGCTTCTCA 1060
 Db 962 ACCCACTGCTCTCTCTCGGAGCAGCC---AGGAGGTGTGACAGAGGCTTCTCA 1021
 Qy 1061 CTCAGAGTGGAAACTCAGATGTGTATC---AATCTCACTTCTGAGACATGACA 1116
 Db 1022 CTCGAGAGGGGAGAACCTCAATGTTCTCCCAAAATTCCTTGTGAGACCTTTACA 1081
 Qy 1117 GCGCTCTGAGGTTTCAAGCCTT 1142
 Db 1082 GCGCTCTGAGGTTTCTCCCAT 1107

RESULT 2
 BX343101 1114 bp mRNA linear EST 02-MAY-2003
 LOCUS BX343101 Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED
 DEFINITION Homo sapiens cDNA clone CS0DL009YD15 5-PRIME, mRNA sequence.
 ACCESSION BX343101
 VERSION BX343101.1 GI:30334170
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1114)
 AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polyes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seque@genoscope.cns.fr, Web: www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 4968.r For
 more information about this cluster, see
 http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CS0DL009YD15&cluster=4968.r. Contact :
 Feng Liang Email: fliang@lifetech.com URL :
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
 Faraday Avenue Genoscope Sequence ID : CS0DL009YD15

FEATURES

source

Location/Qualifiers

1. 1114
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DL009YD15"
 /cell_type="B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"
 /cell_line="RAMOS CELL LINE"
 /clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"

/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
 BASE COUNT 247 a 291 c 360 g 202 t 14 others

ORIGIN

Query Match 52.9%; Score 859; DB 13; Length 1114;
 Best Local Similarity 92.8%; Pred. No. 2.1e-172;
 Matches 966; Conservative 8; Mismatches 60; Indels 7; Gaps 7;

```

Db      72 GTGCGCTCCGACCTCGGCGCTG66CGGCGCGCGGCTG66GAAAGGGCGGCGGCGG
Qy      66 GACCCGATGCGCGGAGCGGAGCGGCGGAGTGGCTTCGCGGAGGCGGAGCACTGCGAG
Db      132 GACCCGATGCGCGGAGCGGAGCGGAGTGGCTTCGCGGAGGCGGAGCACTGCGAG
Qy      126 CCGCGGCGGAGGCGGAGCGGAGTGGCTTCGCGGAGGCGGAGCACTGCGAG
Db      191 CCGCGGCGGAGGCGGAGTGGCTTCGCGGAGGCGGAGCACTGCGAG
Qy      186 CACTGCGGAGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG
Db      250 CACTGCGGAGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG
Qy      246 GGTGGAACAGCGGAGCGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG
Db      310 GGTGGAACAGCGGAGCGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG
Qy      306 GACGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
Db      370 GACGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
Qy      366 CTTGATTAATGATTTGATGACAGAGTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
Db      430 CTTGATTAATGATTTGATGACAGAGTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
Qy      426 GGTGCGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG
Db      490 GGTGCGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG
Qy      486 TGGGAGCGTGGTCTGTTGAGAGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG
Db      550 TGGGAGCGTGGTCTGTTGAGAGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG
Qy      546 GTTCCACTGGGCTCTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
Db      610 GTTCCACTGGGCTCTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
Qy      606 CCGGAGCGTGGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG
Db      670 CCGGAGCGTGGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG
Qy      666 GAAGCGGCGCTTGGAGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG
Db      730 GAAGCGGCGCTTGGAGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG
Qy      726 ACAGAGAGTGGAGATGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG
Db      790 ACAGAGAGTGGAGATGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG
Qy      786 GAATGATGATCTGCAATGAGCAACGAGAGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGG
Db      850 GAATGATGATCTGCAATGAGCAACGAGAGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGG
Qy      846 CTTTTCGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG
Db      909 CTTTTCGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG
Qy      906 GTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
Db      968 GT-CACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
Qy      966 TGTGTGTTACGAGAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
Db      1025 TGTGTGTTACGAGAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
Qy      1026 ACCCGAGGAGAGTGGTGGAG 1046
Db      1085 KTAGMRCGAGCTTTATTCAG 1105

```

```

BX372390/c 941 bp mRNA linear EST 08-MAY-2003
LOCUS      BX372390 Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED
DEFINITION Homo sapiens cDNA clone CS0DL009YD15 3-PRIME, mRNA sequence.
ACCESSION BX372390
VERSION   BX372390.1 GI:30456111
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 941)
AUTHORS   Li, W.B., Gruber, C., Jesse, J., and Polayes, D.
TITLE     Full-length cDNA libraries and normalization
JOURNAL   Unpublished
COMMENT   Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segref@genoscope.cns.fr, Web: www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 4968.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0BA1046ZH05_CS04360_1cluster=4968.r.
Contact : Feng Liang Email : fliang@life.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0BA1046ZH05_CS04360_1.
FEATURES
source
1..941
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DL009YD15"
/cell_line="B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"
/clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE) COT
25-NORMALIZED"
/note="1st strand cDNA was primed with a Notti-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT 196 a 277 c 251 g 217 t
ORIGIN
Query Match 51.8%; Score 841.2; DB 13; Length 941;
Best Local Similarity 96.8%; Pred. No. 1.3e-168;
Matches 912; Conservative 0; Mismatches 23; Indels 7; Gaps 5;
Qy 339 ACAGTACATTTTGAACCATCCAGATGCGCTTGAATATGATTTGATGACAGGACTGTGAA 398
Db 940 ACAGTCAATTTTGAACCAACAAATGCTTGAATATGATTTGAGGACAGGACTTTGAA 881
Qy 399 GAACATGTTGAGAGGCAAAACGGTGA-GGTCCGACATGATTTTGTGACACACTCAA 457
Db 880 AACCATGTTGAGAGGCAAAACGGTGAAGGTGCGCTTATGATTTTGAACCCATCAA 821
Qy 458 GGT--ACCAAGACCAAGTGTCTACCTTGC-GGACTGTTCTGTTGAGGCGATCT 514
Db 820 AGGTACACAGAGACCAAGTGTCTACCTTGCAGGACGTTGTTGAGGCGATCT 761
Qy 515 TGTGTTTACAGCAGAGAGATCCGGG-ACATGTTCACTGCGCTTCTGAGACAC 573
Db 760 TGTGTTTACAGCAGAGAGATCCGGGAAACATGTTCACTGCGCTTCTGAGACAC 701
Qy 574 GACTCCGAGTGAAGTGTCTGAAAGAGTCTCCGAGAGTGCAGCGGAGGAGGAGCTTG 633
Db 700 GACTCCGAGTGAAGTGTCTGAAAGAGTCTCCGAGAGTGCAGCGGAGGAGGAGCTTG 641
Qy 634 GAGCAGATTTGACGAGTACACACCTTGTGTAAGCCGGCTTGTGAGGAGTTTGCCTG 693
Db 640 GAGCAGATTTGACGAGTACACACCTTGTGTAAGCCGGCTTGTGAGGAGTTTGCCTG 581
Qy 694 CCGACAAAGATTAAGCCAGTGTGATCCACGAGAGGTGACATATGTTGCAATC 753

```

Db	Accession	Version	Source	Organism	Reference Authors Title Journal Comment
Db	580				
QY	754				
Db	520				
QY	814				
Db	460				
QY	874				
Db	400				
QY	934				
Db	340				
QY	994				
Db	280				
QY	1054				
Db	220				
QY	1114				
Db	160				
QY	1174				
Db	100				
QY	1234				
Db	40				
RESULT 4					
LOCUS					
DEFINITION					
ACCESSION					
VERSION					
KEYWORDS					
SOURCE					
ORGANISM					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
COMMENT					

```
FEATURES
source
location/Qualifiers
1..916
/organism="Homo sapiens"
/mol_type="rRNA"
/db_xref="taxon:9606"
/cclone="IMAGE:4903819"
```

Query Match	Similarity	Score	DB	Length	916;
Beat Local	97.8%	Pred. No. 2.3e-164;			
Matches	896;	Conservative	0;	Mismatches	13;
				Indels	7;
				Gaps	6;
BASE COUNT	203 a	233 c	309 g	165 t	
ORIGIN	<p>/tissue_type="rhabdomyosarcoma" /lab_host="DH10B (phage-resistant)" /clone_id="NH1_MGC_17" /note="Organ: muscle; Vector: pOTB7; Site_1: EcoRI; Site_2: XhoI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAGAG(G). Size-selected for average insert size 1.8kb. Library constructed by Ling Hong in the Laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."</p>				
15	GACCTCGGCGCTGGGCGGGCGCCGCCGGGAGAGGGCGCGGGAGATCCCATG	74			
1	GACCTCGGCGCTGGGCGGGCGCCGCCGGGAGAGGGCGCGGGAGATCCCATG	60			
75	CGCGGAGCGGAGGCGGAGATGGCTGGGCGGGAGGCGAACAATGCGAGACCCCGCGCC	134			
61	CGCGGAGCGGAGGCGGAGATGGCTGGGCGGGAGGCGAACAATGCGAGACCCCGCGCC	120			
135	GGAGGCCACCGTCCGCACACGCGGCCCTTCTGATAGGGGTGAGCGCGGCACCTGCCAG	194			
121	GAGGCGGACCGTCCGCACACGCGGCCCTTCTGATAGGGGTGAGCGCGGCACCTGCCAG	180			
195	CGGGAAGTCACCGTGTGTGAGAAATATATGAGATTGCTGGGACAGAACGAGGTGGAACA	254			
181	CGGGAAGTCACCGTGTGTGAGAAATATATGAGATTGCTGGGACAGAACGAGGTGGAACA	240			
255	GCGGACGCGAAGGCGTCACTCGAGCGAGACAGGTTCTACAAGGCTCTGACGGCGAGA	314			
241	GCGGACGCGAAGGCGTCACTCGAGCGAGACAGGTTCTACAAGGCTCTGACGGCGAGA	300			
315	GCAGAGGCCCAAGGCTTGAAGAAGACAGTACATTTTGGACCATCCAGATGCCCTTGATTA	374			
301	GCAGAGGCCCAAGGCTTGAAGAAGACAGTACATTTTGGACCATCCAGATGCCCTTGATTA	360			
375	TGATTTGATGACACGAGACTCTGAAGAACATCTGTGAGGGCAAAACGTTGAGGTGCCAC	434			
361	TGATTTGATGACACGAGACTCTGAAGAACATCTGTGAGGGCAAAACGTTGAGGTGCCAC	420			
435	CTATGATTTTGTGACACATCAAGGTTACAGACACACGGTGGCTTACCTCGGGAGAGT	494			
421	CTATGATTTTGTGACACATCAAGGTTACAGACACACGGTGGCTTACCTCGGGAGAGT	480			
495	GATTCTGTTTGAAGGACATTTGTGTTCTACAGCAGAGAGATCCGGGACATGTTCAACT	554			
481	GATTCTGTTTGAAGGACATTTGTGTTCTACAGCAGAGAGATCCGGGACATGTTCAACT	540			
555	GCGCCTCTTGTGACACACGACTCCGACGTCAGGCTGTCTCGAAGAGTTCTCCGGAGCT	614			
541	GCGCCTCTTGTGACACACGACTCCGACGTCAGGCTGTCTCGAAGAGTTCTCCGGAGCT	600			
615	GCGCGGAGGAGGAGACCTGAGAGAGATTTGACGCAATGACCAACTCTTCGGAAGCGCGC	674			
601	GCGCGGAGGAGGAGACCTGAGAGAGATTTGACGCAATGACCAACTCTTCGGAAGCGCGC	660			
675	CTTGAGAGAGTTCTGCGCTCGGACAAAGAGATATGCGCATGTGATCAATCCACAGAGAGT	734			
661	CTTGAGAGAGTTCTGCGCTCGGACAAAGAGATATGCGCATGTGATCAATCCACAGAGAGT	720			
735	GGACAAATA-TGGTTGCCATCAACTGATCTGACAGACA-TCCAGACATTC-TGAAATGG	791			
721	GGACAAATAATGTTGGTTCATCAACTGATCTGACAGACATTCAGAGACATTCCTGAATGG	780			
792	TGACATCTGCAAAATGGCACCGAGAGAGGTCCAAAT-GGGCGGAGCTACAG--CGAACCTT	848			
781	TGACATCTGCAAAATGGCACCGAGAGAGGTCCAAATGAGGGCGGAGGCTACCAAGGCGGACCTTT	840			

QY 849 TTCTGAGCCGAGGGAGACCCCTGGG-ATGCTGACCTCTGCAAAAGGTCAATTTGGAGT 907
 DB 841 TTCTGAGACGAGGGAGAACCCCTGGAGATGCTGACCTTCTGAAACGGTCCCATTTGGAGT 900

QY 908 CCAGCAGCAGACCCCA 923
 DB 901 CCAGCAGCAGACCCCA 916

RESULT 5
 BX400889
 LOCUS 998 bp mRNA linear EST 13-MAY-2003
 DEFINITION BX400889 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
 CDNA clone CS0DK003YL24 5-PRIME, mRNA sequence.
 ACCESSION BX400889
 VERSION BX400889.1 GI:30622359
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 998)
 Title: Full-length cDNA libraries and normalization
 JOURNAL Unpublished
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by life technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 4968.r For
 more information about this cluster, see
 http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CS0DK003DF120P1&cluster=4968.r. Contact :
 Feng Liang Email : fliang@lifetech.com URL :
 http://fulllength.invitrogen.com/Invitrogen Corporation 1600
 Paraday Avenue Genoscope sequence ID : CS0DK003DF120P1.

FEATURES
 Source Location/Qualifiers
 1..998
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DK003YL24"
 /cell_type="HELA CELLS COT 25-NORMALIZED"
 /cell_line="HELA"
 /clone_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"
 /note="1st strand cDNA was primed with a NciI-oligo (dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and cloned into the Not I and EcoR V
 sites of the pCMVSPORT 6 vector. Library was normalized."

BASE COUNT 223 a 263 c 314 g 186 t 12 others

ORIGIN
 Query Match 50.6%; Score 821; DB 13; Length 998;
 Best Local Similarity 96.1%; Pred. No. 2.5e-164;
 Matches 900; Conservative 4; Mismatches 26; Indels 7; Gaps 6;

QY 84 GAGAGCCGAGATGAGCTTGGCGGAGGCGAAGACTGGAGAGCCCGCGGAGAGCGCGA 143
 DB 57 GAGATGCCGAGATGAGCTTGGCGGAGGCGAAGACTGGAGAGCCCGCGGAGAGCGCGA 115

QY 144 CCGTCCGACACCGCGCCCTTCTGATGAGGGGTGAGGGGCGACCTGCCAGCGGAAATC 203
 DB 116 CCGTCCGACACCGCGCCCTTCTGATGAGGGGTGAGGGGCGACCTGCCAGCGGAAATC 175

QY 204 GACCGTGTGAGAAAGATCATGAGTGTGAGCAGAGCAGAGGTGAGACAGCGCGACGC 263
 DB 176 GACCGTGTGAGAAAGATCATGAGTGTGAGCAGAGCAGAGGTGAGACAGCGCGACGC 235

QY 264 GAAAGTGTGATCTCTGAGCAGAGCAGGTTCTCAAGGTCTTGAAGCGCAGAGCAGAAAGC 323
 DB 236 GAAAGTGTGATCTCTGAGCAGAGCAGGTTCTCAAGGTCTTGAAGCGCAGAGCAGAAAGC 295

QY 324 CAAGGCTTGAAGAGCAGTACAAATTTTGACCATCCAGATGCGCTTGTATATGATTTGAT 383
 DB 296 CAAGGCTTGAAGAGCAGTACAAATTTTGACCATCCAGATGCGCTTGTATATGATTTGAT 355

QY 384 GCACAGACTCTGAAGACATGCTGAGAGGCAAAAAGGTGAGGTGCGGACCTATGATTT 443
 DB 356 GCACAGACTCTGAAGACATGCTGAGAGGCAAAAAGGTGAGGTGCGGACCTATGATTT 415

QY 444 TGTGACACTCAAGTTTACAGAGACCAAGGTGTTTACCTGCGGACGTGTTCTGTT 503
 DB 416 TGTGACACTCAAGTTTACAGAGACCAAGGTGTTTACCTGCGGACGTGTTCTGTT 475

QY 504 TGAAGGACATCTTGTGTTTCTTACAGCCAGAGAGATCCGGGACATGTTCCACTGGGCTCTT 563
 DB 476 TGAAGGACATCTTGTGTTTCTTACAGCCAGAGAGATCCGGGACATGTTCCACTGGGCTCTT 535

QY 564 CGTGAACACCGACTCCGAGCTGCTGAGAGTGTCTGAAAGTTCTCCGGAGGTGCGCGAGG 623
 DB 536 CGTGAACACCGACTCCGAGCTGCTGAGAGTGTCTGAAAGTTCTCCGGAGGTGCGCGAGG 595

QY 624 GAGGACCTGAGACAGATTCTGACGACATCACACCTTCTGTAAGCCGCGCTTGAAGA 683
 DB 596 GAGGACCTGAGACAGATTCTGACGACATCACACCTTCTGTAAGCCGCGCTTGAAGA 654

QY 684 GTTCTGCTGCGGACGACAAAGATATGCCATGATCATCCACGAGAGGTGACAAATAT 743
 DB 655 GTTCTGCTGCGGACGACAAAGATATGCCATGATCATCCACGAGAGGTGACAAATAT 714

QY 744 GGTTCATCAACCTGATGCTGAGACGACATCCAGACATTTGTAATGTAATCTGCA 803
 DB 715 GGTTCATCAACCTGATGCTGAGACGACATCCAGACATTTGTAATGTAATCTGCA 774

QY 804 ATGGACCCGAGAGGTTCCATGAGGCTTCAATGAGCTTCAAGCCGACCTTTTCTGAGCAGGGGA 863
 DB 775 ATGGACCCGAGAGGTTCCATGAGGCTTCAATGAGGCTTCAAGCCGACCTTTTCTGAGCAGGGGA 833

QY 864 CCAACCTGGGAGTCTGACCTTGGGCAAGGTCACATTTGGAGTCCAGACGACGACCCCA 923
 DB 834 CCAACCTGGGAGTCTGACCTTGGGCAAGGTCACATTTGGAGTCCAGACGACGACCCCA 892

QY 924 CTGAGGAGGCTGCGAGAGCTCAGGAGAGGTCTCCCGCGGACATGTTCTGAGGAGCTGA 983
 DB 893 CTGAGGAGGCTGCGAGAGCTCAGGAGAGGTCTCCCGCGGACATGTTCTGAGGAGGAG-75 949

QY 984 GCCTGGGAGCGCCACCAACCACTGTTCTCTC 1020
 DB 950 AGCTGGGAGCGCCACCAACCACTGTTCTCTC 986

RESULT 6
 BG491358 860 bp mRNA linear EST 27-MAR-2001
 LOCUS 602535642F1 NIH_MGC_41 Homo sapiens CDNA clone IMAGE:4684738 5',
 DEFINITION mRNA sequence.
 ACCESSION BG491358
 VERSION BG491358.1 GI:13452870
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 860)
 Title: NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaabs-remail.nih.gov
 Tissue Procurement: DCTP/DRP
 CDNA Library Preparation: Ling Hong/Rubin Laboratory
 DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
 Clone distribution: MGC clone distribution information can be


```

Db      185 CAGCGGAGAGTCGACCGTGTGTGAGAGATCATGAGTTGCTGGACAGACGAGGTGGA 244
Qy      252 ACAGCGGACACGGAAGGTGTCATCTCGAGCCAGAGACAGTTCTTACAAAGTCTCTGAGCGC 311
Db      245 ACAGCGGACACGGAAGGTGTCATCTCGAGCCAGAGACAGTTCTTACAAAGTCTCTGAGCGC 304
Qy      312 AGAGCAGAGAGCCCAAGGCTTTGAAAGAGACAGTACATTTTGGACCATCCAGATGCTTTGA 371
Db      305 AGAGCAGAGAGCCCAAGGCTTTGAAAGAGACAGTACATTTTGGACCATCCAGATGCTTTGA 364
Qy      372 TAATGATTTGATGACAGACCTCTGAAAGAACATGCTGAGAGGCGCAAAACGCTGAGAGTCC 431
Db      365 TAATGATTTGATGACAGACCTCTGAAAGAACATGCTGAGAGGCGCAAAACGCTGAGAGTCC 424
Qy      432 GACCTATGATTTTGTGACACACTCAAGTTACCAAGAACCAAGGCTGCTTACCTTGGGA 491
Db      425 GACCTATGATTTTGTGACACACTCAAGTTACCAAGAACCAAGGCTGCTTACCTTGGGA 484
Qy      492 CGTGGTCTGTGTTGAGGAGCATCTTGTGTTCTTACAGCAGAGAGATCCGGGACATGTTCCA 551
Db      485 CGTGGTCTGTGTTGAGGAGCATCTTGTGTTCTTACAGCAGAGAGATCCGGGACATGTTCCA 544
Qy      552 CTTGCGGCTCTTCTGTGACACCGACCTCCAGCTGAGGCTGTCTTGAAGAGTTCTCCGGGA 611
Db      545 CTTGCGGCTCTTCTGTGACACCGACCTCCAGCTGAGGCTGTCTTGAAGAGTTCTCCGGGA 604
Qy      612 CTTGCGGCTCTTCTGTGACACCGACCTCTGAGCAGATTTCTGACGATCACCACTTCTGGAAGCC 671
Db      605 CTTGCGGCTCTTCTGTGACACCGACCTCTGAGCAGATTTCTGACGATCACCACTTCTGGAAGCC 664
Qy      672 GGCCTTGAAGAGATTTCTGCTGCGGACCAAGAAATATGCGAGTGTATCATCCACAGAG 731
Db      665 GGCCTTGAAGAGATTTCTGCTGCGGACCAAGAAATATGCGAGTGTATCATCCACAGAG 724
Qy      732 AGTGAACAATATGTTGTCATCAACTGATCTGTGACAGACATCCAGAGACATTTGAAATGG 791
Db      725 AGTGAACAATATGTTGTCATCAACTGATCTGTGACAGACATCCAGAGACATTTGAAATGG 784
Qy      792 TGACATCTGCAAAATGACACCGAGAGGCTCCATGCGGCGAGCTTACAAAGCGGACCTTTTC 851
Db      785 TGACATCTGCAAAATGACACCGAGAGGCTCCATGCGGCGAG-CTAACAAAGCGGACCTTTTC 843
Qy      852 TGAGCCAGGAGGACCACTCGGAGATGTCGACCTCTGGGAAAACGCTCAATTTGAGATCCAG 911
Db      844 TGAGCCAGGAGGAC-CACTCGGAGATGTCGACCTCTGGGAAAACGCTCAATTTGAGAT-CAAG 901
Qy      912 CAGCAGACCCCACTGAGAGGCTGCGGAGCCTTCAAGGAGAGTCTCCCGCCGCGCATGTGTG 971
Db      902 CAGCAGAC-CCCACTGAGAGGCTGCG- -GAGCTCAGGCAAGGTCTCC- -GCCGGCATGTGTG 955
Qy      972 TTGAGGAGCTGAGAGCTTGGGAGCGCCACCCACACCCCATGCTTCTTGGCGGACCCCA 1031
Db      956 TTGAG- -GAGCTGAGAGCTTGGGAGCG- - -ACGCCACACACACCACTGCTTCTTGGCGGACCCCA 1007
Qy      1032 GGGGAGTGTAGACGAGGAGGCTTCTTCACTCACTCAAGGA 1067
Db      1008 AGGAGATGTAGAGGAGGCTTCTTMTMTAGAGTGGGA 1043

```

```

RESULT 8
LOCUS    BM803129
DEFINITION
  BM803129 1039 bp mRNA linear EST 05-MAR-2002
  AGENCOURT_6468837 NIH_MGC_125 Homo sapiens cDNA clone IMAGE:5721513
  5', mRNA sequence.
ACCESSION
  BM803129
VERSION  BM803129.1 GI:19119952
KEYWORDS
  EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
  1 (bases 1 to 1039)
  NIH-MGC http://mgs.nci.nih.gov/.

```

```

TITLE    National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL  Unpublished
COMMENT  Contact: Robert Strausberg, Ph.D.
  Email: cgepds-remail.nih.gov
  Tissue Procurement: Invitrogen
  cDNA Library Preparation: Life Technologies, Inc.
  DNA Sequencing by: Agencourt Bioscience Corporation
  Clone distribution: MGC clone distribution information can be
  found through the I.M.A.G.E. Consortium/LLNL at:
  http://image.llnl.gov
  Plate: LLNL12706 row: b column: 10
  High quality sequence stop: 652.
FEATURES
  source
    1..1039
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /clone="IMAGE:5721513"
    /lab_host="DH10B"
    /clone_lib="NIH_MGC_125"
    /note="Organ: ovary (pool of 3); Vector: pCMV-Sport6;
    Site 1: EcoRV (destroyed); Site 2: NotI; RNA source pool
    of three ovaries, from females ranging in age from 38 to
    49 yo. Library is oligo-dT primed and directionally cloned
    (EcoRV site is destroyed upon cloning). Average insert
    size 2.1 kb, insert size range 1-3.5 kb. Library is
    normalized and enriched for full-length clones and was
    constructed by C. Gruber (Invitrogen). Research Genetics
    tracking code 036."
BASE COUNT  240 a 266 c 309 g 220 t 4 others
ORIGIN
Query Match 48.3%; Score 783.6; DB 12; Length 1039;
Best Local Similarity 98.3%; Pred. No. 2.3e-156;
Matches 812; Conservative 0; Mismatches 11; Indels 3; Gaps 2;
Qy      323 CCAAGGCTTTGAAAGAGACAGTACATTTTGGACCATCCAGATGCTTTGATTAAGATTTGA 382
Db      1 CCAAGGCTTTGAAAGAGACAGTACATTTTGGACCATCCAGATGCTTTGATTAAGATTTGA 60
Qy      383 TGACAGAGACTCTGAGAGACATGTTGAGGAGGCAAAACGCTGAGAGTGCAGACCTATGATT 442
Db      61 TGACAGAGACTCTGAGAGACATGTTGAGGAGGCAAAACGCTGAGAGTGCAGACCTATGATT 120
Qy      443 TTGTGACACACTCAAGTTTACAGAGACCAAGGTGTTTACCTTGGGACGTTGTTCTGT 502
Db      121 TTGTGACACACTCAAGTTTACAGAGACCAAGGTGTTTACCTTGGGACGTTGTTCTGT 180
Qy      503 TTGAGGAGCATCTTGTGTTCTTACAGCAGAGAGATCCGGGACATGTTCCACCTGGCCTCT 562
Db      181 TTGAGGAGCATCTTGTGTTCTTACAGCAGAGAGATCCGGGACATGTTCCACCTGGCCTCT 240
Qy      563 TCGTGAACACCGACTCCAGAGTCAAGGCTGTCTTGAAGAGTCTCCGGGAGTGCAGAG 622
Db      241 TCGTGAACACCGACTCCAGAGTCAAGGCTGTCTTGAAGAGTCTCCGGGAGTGCAGAG 300
Qy      623 GAGAGGAGCTGAGAGAGATTTGAGCAGAGTACACCACTTGTGAAGCCGGCTTGAAG 682
Db      301 GAGAGGAGCTGAGAGAGATTTGAGCAGAGTACACCACTTGTGAAGCCGGCTTGAAG 360
Qy      683 AGTTTCTGCTGCGGACAAAGAGATGATCCGAGTATATATCCACAGAGAGTGAACAAAT 742
Db      361 AGTTTCTGCTGCGGACAAAGAGATGATCCGAGTATATATCCGAGAGTGAACAAAT 420
Qy      743 TGGTTCACATCAACTGATCTGTGACAGCATTCAGAGACATTTGATGATGATCATCTGCA 802
Db      421 TGGTTCACATCAACTGATCTGTGACAGCATTCAGAGACATTTGATGATGATCATCTGCA 480
Qy      803 AATGACACGAGAGGAGGTCATGAGGCGGAGCTTCAAGCGGACCTTTTGTGAGCCAGGG 862
Db      481 AATGACACGAGAGGAGGTCATGAGGCGGAGCTTCAAGCGGACCTTTTGTGAGCCAGGG 540

```

QY 863 ACCACCTGGAGTGTGACTCTGTGGCAACCGGTACATTTGGAGTCCAGACAGACCC 922
 Db 541 ACCACCTGGAGTGTGACTCTGTGGCAACCGGTACATTTGGAGTCCAGACAGACCC 600
 QY 923 ACTGAGGGGCTGCGGAGCTTCAAGGAGAGTCTCCGCGCGAGATGTGTTCAGGAGCTG 982
 Db 601 ACTGAGGGGCTGCGGAGCTTCAAGGAGAGTCTCCGCGCGAGATGTGTTCAGGAGCTG 660
 QY 983 AGCCTGGGAGCGCCACCCACCAACCACTGCTTCCCTGTGGCGACCCCGAGGAGTGTTA 1042
 Db 661 AGCCTGGGAGCGCCACCCACCAACCACTGCTTCCCTGTGGCGACCCCGAGGAGTGTTA 720
 QY 1043 GCACGAGGCGCTTCTCTCACTGAGAGTGG-AACTCAGATGTGTCACTCAGACTCAATT 1101
 Db 721 GCACGAGGCGCTTCTCTCACTGAGAGTGGAAACTCAGATGTGTCACTCAGACTCAATT 780
 QY 1102 GCTGGGACACTGACAGGCG--TTCTGAGGTTTTCAGGCACTTAA 1145
 Db 781 GCTGGGACACTGACAGGCGGTTCTTGAAGGTTTTCAGGCACTTAA 826

RESULT 9
 BG491384 869 bp mRNA linear EST 27-MAR-2001
 LOCUS 602535670P1 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:4684936 5',
 DEFINITION mRNA sequence.
 ACCESSION BG491384
 VERSION BG491384.1 GI:13452896
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 869)
 NIH-MGC http://mgi.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cga@bbs-remail.nih.gov
 Tissue Procurement: DCTD/DRP
 CDNA Library Preparation: Ling Hong/Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: LICM1493 row: k column: 17
 High quality sequence stop: 863.
 Location/Qualifiers
 1. 869
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4684936"
 /label_type="amelanotic melanoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_id="NIH_MGC_41"
 /note="Organ: skin; Vector: pOTB7; Site 1: XhoI; Site 2:
 EcoRI; cDNA made by oligo-dT priming. Directionally cloned
 into EORI/XhoI sites using the following 5' adaptor:
 GGCAAGAG(G). Library constructed by Ling Hong in the
 laboratory of Gerald M. Rubin (University of California,
 Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and
 Superscript II RT (Life Technologies). Note: this is a
 NIH_MGC Library."
 BASE COUNT 197 a 223 c 291 g 156 t 2 others
 ORIGIN
 Query Match 48.1%; Score 781; DB 10; Length 869;
 Best Local Similarity 97.0%; Pred. No. 8.1e-156;
 Matches 837; Conservative 0; Mismatches 21; Indels 5; Gaps 4;
 QY 44 CCGGGAAGGGCGGGCGCGGGAACCGATGCGGGAGCGGAGCCGAGATGGCTTCGG 103

Db 2 CTGGGGAAGGGCGGGCGCGGGAACCGATGCGGGAGCGGAGCCGAGATGGCTTCGG 61
 QY 104 CCGGAGGCGGAAGACTGCGAGAGACCCCGCGCGGAGCCGATCCGACACAGGCGCTT 163
 Db 62 CCGGAGGCGGAAGACTGCGAGAGCCCGCGCGGAGCCGATCCGACACAGGCGCTT 121
 QY 164 TCTGATAGGGGTGAGCGCGGCACTGCCAGCGGGAAGTGCACCGTGTGGAAGATCA 223
 Db 122 TCTGATAGGGGTGAGCGCGGCACTGCCAGCGGGAAGTGCACCGTGTGGAAGATCA 181
 QY 224 TGAAGTTGCTGGGACAGAAAGAGGTGAAACAGCGGAGCGGAAGGTGTATCTGAGCC 283
 Db 182 TGAAGTTGCTGGGACAGAAAGAGGTGAAACAGCGGAGCGGAAGGTGTATCTGAGCC 241
 QY 284 AGGACAGGTTCTACAAAGTCTCTGACCGGAGAGAGAAAGCCCTTGAAGACACT 343
 Db 242 AGGACAGGTTCTACAAAGTCTCTGACCGGAGAGAGAAAGCCCTTGAAGACAGT 301
 QY 344 ACAATTTTGACATCCAGATGCTTTGATATATGATGATGACAGGACTCTGAAGACA 403
 Db 302 ACAATTTTGACATCCAGATGCTTTGATATATGATGATGACAGGACTCTGAAGACA 361
 QY 404 TCGTGAAGGCAAAACGGTGAAGGTGCGGACCTATGATTTTGTGACACTCAAGGTTAC 463
 Db 362 TCGTGAAGGCAAAACGGTGAAGGTGCGGACCTATGATTTTGTGACACTCAAGGTTAC 421
 QY 464 CAGAGACCAAGGTGTCTACCTCGCGGAGCGTGTCTGTGTTGAGGCACTTGTGTCT 523
 Db 422 CAGAGACCAAGGTGTCTACCTCGCGGAGCGTGTCTGTGTTGAGGCACTTGTGTCT 481
 QY 524 ACAGCCAGAGATCCGGGACATGTTCCACTCGGCGCTTCTGTGAGACACCGACTCCGAG 583
 Db 482 ACAGCCAGAGATCCGGGACATGTTCCACTCGGCGCTTCTGTGAGACACCGACTCCGAG 541
 QY 584 TCAAGCTGTCTGAAGAAGTTCCTCGGAGCGTGGCGGAGGAGGACTTGGAGCAATTC 643
 Db 542 TCAAGCTGTCTGAAGAAGTTCCTCGGAGCGTGGCGGAGGAGGACTTGGAGCAATTC 601
 QY 644 TGAAGCAGTACACACCTTCGTGAAGCCGCGCTTGTGAGAGTTCGTCGCAACAAAG- 702
 Db 602 TGAAGCAGTACACACCTTCGTGAAGCCGCGCTTGTGAGAGTTCGTCGCAACAAAG 661
 QY 703 AAGTATGCCGATGTGATCCACGAGAGTGAACATAT-GGTTGCCATCAACTGAT 761
 Db 662 AAGTATGCCGATGTGATCCACGAGAGTGAACATATGGTTGCCATCAACTGAT 721
 QY 762 CGTGACGACATCCAGACATTTCTGAATGTGACATCTGCMAATGGCA-CCGAGAGGGT 820
 Db 722 CGTGACGACATCCAGACATTTCTGAATGTGACATCTGCMAATGGCACCCGAGAGGGT 781
 QY 821 CCAATGGGGGAGACTCAAGCGGACCTTTTCTGAGCGAGGGGACCACTCGGATGCTGA 880
 Db 782 CCAATGGGGGAGACTCAAGCGGACCTTTTCTGAGCGAGGGG--ACCACTGGGATGCTGA 839
 QY 881 CCTCTGGCAAAACGGTCAATTTG 903
 Db 840 CCTCTGGCAAAACGGTCAATTTG 862

RESULT 10
 BG492079 929 bp mRNA linear EST 27-MAR-2001
 LOCUS 602555670T1 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:4684936 3',
 DEFINITION mRNA sequence.
 ACCESSION BG492079
 VERSION BG492079.1 GI:13453591
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 929)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabs-remail.nih.gov
 Tissue Procurement: DCTD/DTF
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
 Plate: L1CM1493 row: k column: 17
 High quality sequence start: 24
 High quality sequence stop: 820.
 Location/Qualifiers

FEATURES

source

1. 929
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4684936"
 /tissue_type="amelanotic melanoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_id="NIH MGC 41"
 /note="Organ: skin; Vector: pOTB7; Site 1: XhoI; Site 2:
 EcoRI; cDNA made by oligo-dT priming. Directionally cloned
 into EcoRI/XhoI sites using the following 5' adaptor:
 GGCGAGG(G). Library constructed by Ling Hong in the
 Laboratory of Gerald M. Rubin (University of California,
 Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
 Superscript II RT (Life Technologies). Note: this is a
 NIH MGC Library."

BASE COUNT 195 a 280 c 249 g 205 t
 ORIGIN

Query Match 47.8%; Score 776.6; DB 10; Length 929;
 Best Local Similarity 96.0%; Pred. No. 7.1e-155;
 Matches 871; Conservative 0; Mismatches 29; Indels 7; Gaps 7;

QY 443 TTGTGACACACTCAAGGTTACAGAGACACGAGTGTCTACCTTCGCGAGCTGTTCTGT 502
 DB 912 TTGGGAACAACCTCAAGTTACAGAGACCCGCGTGTCTACCTTCGCGAGCTGTTCTGT 853
 QY 503 TTGAGGCACTTGTGTTCTGAC-GCAGAGATCCGGG-ACATGTTCCACCTGGCCCT 560
 DB 852 TCGAGGCAATTTGTGTCTGACGCGCAGAGATCCGGGAACATGTTCCACCTGGCCCT 793
 QY 561 CTTCGTGG-ACACGACTCCGAC-GTCAGGCTGTCTCGAAG-AGTTCTCCGAGACGTG-C 616
 DB 792 CTTGTGGAACACCGAATCCGACTGTCAAGCTGTCTGAAATGTTCTCCGAGACGTGTC 733
 QY 617 GCCGAGGAGGAGCTTGAGAGCATTTCTGACGAGTACACCACTTCTGTGAAGCCGCT 676
 DB 732 GCCGAGGAGGAGCTTGAGAGCATTTCTGACGAGTACACCACTTCTGTGAAGCCGCT 673
 QY 677 TCGAGAGTTCTGCTGCTCCGACAAAGATATGCCATGTGATCATCCACGAGAGATTGG 736
 DB 672 TCGAGAGTTCTGCTGCTCCGACAAAGATATGCCATGTGATCATCCACGAGAGATTGG 613
 QY 737 ACAATATGTTGCTCATCACTGATCGTGCAGCATTCAGAGCATTTCTGAATGTGACA 796
 DB 612 ACAATATGTTGCTCATCACTGATCGTGCAGCATTCAGAGCATTTCTGAATGTGACA 553
 QY 797 TCTGCAATGAGCAGCAGAGAGGTTCCAAATGGGCGAGCTACAGCGACCTT-TTCTGAG 855
 DB 552 TCTGCAATGAGCAGCAGAGAGGTTCCAAATGGGCGAGCTACAGCGACCTTGTCTGAG 493
 QY 856 CCGAGGAGACACCTTGAGATGCTGACCTTGGGCAAAAGTCCAACTTGGAGTCCAGAGC 915
 DB 492 CCGAGGAGACACCTTGAGATGCTGACCTTGGGCAAAAGTCCAACTTGGAGTCCAGAGC 433
 QY 916 AGACCCCACTGAGGAGGCTGCGAGACCTCAGGGCAGGTCTCCGCGCCGAGCATGTGTCTCA 975

DB 432 AGACCCCACTGAGGAGGCTGCGAGACCTCAGGGCAGGTCTCCGCGCCGAGCATGTGTCTCA 373
 QY 976 GGAATGAGCTTGGGAGACGGCCACACACCCACTGTTCTTCTGGCGGACCCAGGGG 1035
 DB 372 GGAATGAGCTTGGGAGACGGCCACACACCCACTGTTCTTCTGGCGGACCCAGGGG 313
 QY 1036 AGTGTGAGCAGCAGGCTTCTCTCACTCAGAGATGGAATACTCAGATGTGCACTAGACT 1095
 DB 312 AGTGTGAGCAGCAGGCTTCTCTCACTCAGAGATGGAATACTCAGATGTGCACTAGACT 253
 QY 1096 CAATCTGCTGGACACTGACAGGCGTTTCTGAGTTTTCAGCCACTTGAAGCTGTTGCGG 1155
 DB 252 CAATCTGCTGGACACTGACAGGCGTTTCTGAGTTTTCAGCCACTTGAAGCTGTTGCGG 193
 QY 1156 TTTAAAGATCCCTTCTGAGTCACTGAGAAAAGCCAGAAATGTCAGAAACCTGGAGGC 1215
 DB 192 TTTAAAGATCCCTTCTGAGTCACTGAGAAAAGCCAGAAATGTCAGAAACCTGGAGGC 133
 QY 1216 TTCTGTGAGAAATGTGAGCACAATTATTTGGGAAATTTGAGAGACAGCTAGACACTGCG 1275
 DB 132 TTCTGTGAGAAATGTGAGCACAATTATTTGGGAAATTTGAGAGACAGCTAGACACTGCG 73
 QY 1276 TGACCTGATGTTTGTGTAACAGTGAACCCACAGTGGAGAGAGTTTTTCAGTCTGATC 1335
 DB 72 TGACCTGATGTTTGTGTAACAGTGAACCCACAGTGGAGAGAGTTTTTCATACCGCGC 13
 QY 1336 TGCTTCT 1342
 DB 12 GGGTTCT 6

RESULT 11

BO072501

LOCUS 1036 bp mRNA linear EST 02-APR-2002
 DEFINITION AGENCOURT 6838909 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:5761694

ACCESSION BO072501
 VERSION BO072501.1 GI:19901547

KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1036)
 NIH-MGC <http://mgc.nci.nih.gov/>.
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabs-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
 Plate: L1AM12810 row: 1 column: 15
 High quality sequence stop: 644.
 Location/Qualifiers

FEATURES

source

1. 1036
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5761694"
 /lab_host="DH10B"
 /clone_id="NIH MGC 122"
 /note="Organ: pooled lung and spleen; Vector: pCMV-Sport6;
 Site 1: NotI; Site 2: EcoRV (destroyed); RNA source
 anonymous pool of 24 week female lung, 16 week female
 spleen, and 20-22 week male spleens. Library is oligo-dT
 primed and directionally cloned (EcoRV site is destroyed
 upon cloning). Average insert size 1.4 kb, insert size
 range 1-3 kb. Library is normalized and enriched for

Db 360 ATGCCAGATGTCAGAGAGCTGGAGGCTTCTGTGAGATGTGAGCATTATT 419
 Qy 1244 GGGGAATTAGAGAGAGCTTGAACA CTGGCTGGCTGATGTTTGTGACAGTGAAC 1303
 Db 420 GGGGAATTAGAGAGAGCTTGAACA CTGGCTGGCTGATGTTTGTGACAGTGAAC 479
 Qy 1304 CACAGTGGAGAGAGTCTTTCAGTCTGATCTGTTCTTACACATCTACACATTAAT 1363
 Db 480 CACAGTGGAGAGAGTCTTTCAGTCTGATCTGTTCTTACACATCTACACATTAAT 539
 Qy 1364 CAAAGTTTGTGAACAAGTACTTCTTTTTCATGTTACATGTCCTCATGTTTCTG 1423
 Db 540 CAAAGTTTGTGAACAAGTACTTCTTTTTCATGTTACATGTCCTCATGTTTCTG 599
 Qy 1424 TTTTCTGTTTCAATACACAAGGCTGTGTGGCTTACAACTTAATTTCAAGCCAGTG 1483
 Db 600 TTTTCTGTTTCAATACACAAGGCTGTGTGGCTTACAACTTAATTTCAAGCCAGTG 659
 Qy 1484 GTTTCAGTCCAGGCTGGCTTACAGGATATGGGAGCCACTGAGGATGTTTCCCCC 1543
 Db 660 GTTTCAGTCCAGGCTGGCTTACAGGATATGGGAGCCACTGAGGATGTTTCCCCC 719
 Qy 1544 TTGCTGTGCTTAAAGCAGAGAGCGAGCGGATGCCCTGAAGCACCAGCATCA 1603
 Db 720 CTGCTTGTGCTTAAAGCAGAGAGCGAGCGGATGCCCTTGAAGCACCAGCATCA 779
 Qy 1604 CCCAGGCTTGTGCGGCGCCAG 1624
 Db 780 CCCAGGCTTGTGCGGCGCCAG 800

RESULT 13 B0719741

LOCUS B0719741 956 bp mRNA linear EST 16-JUL-2002
 DEFINITION A58NCOURT_8229663 Lupski_dorsal_root_ganglion Homo sapiens cDNA
 clone IMAGE:6184625 5', mRNA sequence.
 B0719741

ACCESSION B0719741 GI:21858638
 VERSION EST.
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens

Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE NIH-MGC http://mgc.nci.nih.gov/
 1 (bases 1 to 956)
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 TITLE Unpublished
 JOURNAL Contact: Robert Strausberg, Ph.D.
 COMMENT Email: cgabbs-remail.nih.gov
 Tissue Procurement: Dr. James R. Lupski
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)
 Cloning distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LMNL at:
 http://image.jnl.gov
 Plate: LAM13574 row: b column: 18
 High quality sequence stop: 552.

FEATURES source

1. 956
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6184625"
 /sex="male"
 /tissue_type="dorsal root ganglia"
 /dev_stage="adult, 36 yr"
 /lab_host="DH108"
 /clone_lib="Lupski_dorsal_root_ganglion"
 /note="Vector: PCMV-SPORE6 (Life Technologies); Site_1:
 NotI; Site_2: SalI; cDNA made by oligo-dT priming.
 Directionally cloned using the following adaptors:
 5'-TCGACCCAGCGGTCCG-3' and

BASE COUNT ORIGIN

Query Match 46.6%; Score 756.4; DB 13; Length 956;
 Best Local Similarity 95.9%; Pred. No. 1.4e-150;
 Matches 798; Conservative 0; Mismatches 31; Indels 3; Gaps 2;

5'-GACTAGTCTAGATCCGAGCGGCCCTT(15)-3'. Size selected >
 1 kb for average insert length 1.7 kb. This is a primary
 library, non-amplified. Library constructed by Life
 Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
 College of Medicine) and is available through Life
 Technologies."

202 a 272 c 296 g 186 t
 88 GCCGAGATGCTTCGGCGGGAGGCGGAAGCTGGAGAGCCCGCGCGGAGCGGACCGT 147
 Db 1 GCCGAGATGCTTCGGCGGGAGGCGGAAGCTGGAGAGCCCGCGCGGAGCGGACCGT 60
 Qy 148 CCGCACCAGCGGCCCTTCTGATAGGAGGAGCGGCGGCACTGCCAGCGGGAAGTGCACC 207
 Db 61 CCGCACCAGCGGCCCTTCTGATAGGAGGAGCGGCGGCACTGCCAGCGGGAAGTGCACC 120
 Qy 208 GTGTGTGAGAGATCATGAGTGTCTGGGACAGACAGAGTGAACAAGCGGCAAGG 267
 Db 121 GTGTGTGAGAGATCATGAGTGTCTGGGACAGACAGAGTGAACAAGCGGCAAGG 180
 Qy 268 GTGTGATCCTGAGCCAGAGCAGGTTCTCAAGAGTCTCAAGCGGCAAGAGGCCAAG 327
 Db 181 GTGTGATCCTGAGCCAGAGCAGGTTCTCAAGAGTCTCAAGCGGCAAGAGGCCAAG 240
 Qy 328 GCCTTGAAGAGCAGTACATTTTGAACATCCAGATGCTTGTATATGATTTGATGAC 387
 Db 241 GCCTTGAAGAGCAGTACATTTTGAACATCCAGATGCTTGTATATGATTTGATGAC 300
 Qy 388 AGGACTCTGAAGAACTCTGTGAGGCGAAACCGTGAAGTCCGACCTATGATTTTGTG 447
 Db 301 AGGACTCTGAAGAACTCTGTGAGGCGAAACCGTGAAGTCCGACCTATGATTTTGTG 360
 Qy 448 ACACTCTCAAGTTACAGAGACAGAGTGTATCCCTGCGGAGCGTGTCTGTTTGTAG 507
 Db 361 ACACTCTCAAGTTACAGAGACAGAGTGTATCCCTGCGGAGCGTGTCTGTTTGTAG 420
 Qy 508 GGCATCTGTGTGTTCTACAGCCAGAGATCCGGGACATGTTCCAGCTGCGCTCTTG 567
 Db 421 GGCATCTGTGTGTTCTACAGCCAGAGATCCGGGACATGTTCCAGCTGCGCTCTTG 480
 Qy 568 GACACCGACTCCGACGTCAGGCTGTCTGAAGGTTCTCCGAGACGTGCGCGAGGAG 627
 Db 481 GACACCGACTCCGACGTCAGGCTGTCTGAAGGTTCTCCGAGACGTGCGCGAGGAG 540
 Qy 628 GACCTGAGCAGATTTGTAGAGCAGTACACACCTTGTGTGAAGCGGCGCTTCCAGAGTT 687
 Db 541 GACCTGAGCAGATTTGTAGAGCAGTACACACCTTGTGTGAAGCGGCGCTTCCAGAGTT 600
 Qy 688 TGCCTGCGGCAAAAGATATGCGGATGTGATCATCCACAGAGAGTGAACAATATG-7 746
 Db 601 TGCCTGCGGCAAAAGATATGCGGATGTGATCATCCCGAGAGATGAACAATATG-7 660
 Qy 747 TGCATCAACCTGATGTCGAGACATCCAGACATTTGGAATGTGACATTCGAATG 806
 Db 661 TGCATCAACCTGATGTCGAGACATCCAGACATTTGGAATGTGACATTCGAATG 720
 Qy 807 GCACCGAGAGAGTCCAAATGAGGAGGAGTACAAAGCGGACCTTTTGAAGC--AGGAGC 864
 Db 721 GCACCGAGAGAGTCCAAATGAGGAGGAGTCCAAAGCGGACCTTTTGAAGC--AGGAGC 780
 Qy 865 CACCTGAGATGTGACCTCTGCAAAACGTCATTTGAGATCCAGACGA 916
 Db 781 CACCTGAGATGTGACCTTTGACCTTAAGTTCCATTTTGGAGGCGCAGCA 832

RESULT 14
 BM459108
 LOCUS BM459108 1051 bp mRNA linear EST 05-FEB-2002


```

DEFINITION  AGENCOURT_6414543 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5557433
5', mRNA sequence.
ACCESSION   BM459108
VERSION     BM459108.1 GI:18508148
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE   1 (bases 1 to 1051)
AUTHORS     NIH-MGC http://mgc.nci.nih.gov/.
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgabbs-remail.nih.gov
            Tissue Procurement: ATCC/DCTD/DRP
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LNL at:
            http://image.lnl.gov
            Plate: LLM12278 row: m column: 18
            High quality sequence stop: 542.
            Location/Qualifiers
                1..1051
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="IMAGE:5557433"
                /tissue_type="melanotic melanoma"
                /lab_host="DH10B (phage-resistant)"
                /clone_1lb="NIH MGC 72"
                /note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;
                Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
                Average insert size 2 kb. Library constructed by Life
                Technologies."
BASE COUNT  237 a      264 c      312 g      221 t      17 others
ORIGIN
Query Match      46.5%; Score 755.2; DB 12; Length 1051;
Best Local Similarity 97.5%; Pred. No. 2.6e-150;
Matches 788; Conservative 0; Mismatches 18; Indels 2; Gaps 2;

QY 89 CCGAGATGCTTCGCGCGGAGGCGGAAGACCTCGAGAGCCCGCGGAGGCGGACCGTC 148
Db 1 CCGAGATGCTTCGCGCGGAGGCGGAAGACCTCGAGAGCCCGCGGAGGCGGACCGTC 60

QY 149 CGCACACAGCGGCCCTTCTCTGATAGGGGTGAGCGCGGCACTGCCAGCGGAAAGTCGACCG 208
Db 61 CGCACACAGCGGCCCTTCTCTGATAGGGGTGAGCGCGGCACTGCCAGCGGAAAGTCGACCG 120

QY 209 TGTGTGGAAGAATATGATGAGTTTGTCTGGAGACGAACGAGTGGAAACAGCGGACGCGGAAGG 268
Db 121 TGTGTGGAAGAATATGATGAGTTTGTCTGGAGACGAACGAGTGGAAACAGCGGACGCGGAAGG 180

QY 269 TGGTCATCTGAGCGCAGGACAGTCTTCTCAAGGTCCTGACGCGAGAGAGCAAGGCGCAAGG 328
Db 181 TGGTCATCTGAGCGCAGGACAGTCTTCTCAAGGTCCTGACGCGAGAGAGCAAGGCGCAAGG 240

QY 329 CTTTGAAGAAGACAGTACATTTTGAACATCCAGATGCCCTTGTGATGATTTGATGACACA 388
Db 241 TCTTGAAGAAGACAGTACATTTTGAACATCCAGATGCCCTTGTGATGATTTGATGACACA 300

QY 389 GGAATCTGAAGAACAATCGTGAAGGCAAAAACGGTGAAGTCCGACCTATGATTTTGTGA 448
Db 301 GGAATCTGAAGAACAATCGTGAAGGCAAAAACGGTGAAGTCCGACCTATGATTTTGTGA 360

QY 449 CACACTCAAGGTTACAGAGACACAGGTCGTCTACCCCTGCGGACGTGGTTCTGTTTGAAG 508
Db 361 CACACTCAAGGTTACAGAGACACAGGTCGTCTACCCCTGCGGACGTGGTTCTGTTTGAAG 420

QY 509 GCATCTTGTGTGTTCTACAGCAGGAGATCGGGAATGTTCCACCTGCGCTCTTCGTGG 568

```

```

Db 421 GCATCTTGTGTGTTCTACAGCAGGAGATCGGGAACATGTTCCACCTGCGCTCTTCGTGG 480
QY 569 ACAACCGATCCGACAGCTGAGGCTGCTTCGAGAGATTTCCCGGAGCGCGGAGGAGG 628
Db 481 ACAACCGATCCGACAGCTGAGGCTGCTTCGAGAGATTTCCCGGAGCGCGGAGGAGG 540
QY 629 ACCTGAGCAGATTTGACACCACTGACAGCACTTCGTAAGCCGCTTCGAGAGTCT 688
Db 541 ACCTGAGCAGATTTGACACCACTGACAGCACTTCGTAAGTGGTCTTCGAGAGTCT 600
QY 689 GCCTGCGCAAGAAAGATGATCCGATGTGATCATCCACGAGAGTGGACATATGCTTG 748
Db 601 GTCTGCGCAAGAAAGATGATGATGTGATCATCCACGAGAGTGGACATATGCTTG 660
QY 749 CAAATCAACCTGATCGACACATCATCCAGACATTCGTAATGTTGATCTGCAATGCG 808
Db 661 ACAAATCAACCTGATCGACATCATCCAGACATTCGTAATGTTGATCTGCAATGCG 720
QY 809 ACCGAGAGAGGTCCTCAATGAGCGGAGCTACAG-CCGACCTTTCTGAGCCAGGAGCA- 866
Db 721 ACCGTAAGAGGTCCTCAATGAGCGGAGCTACAGTCCGACCTTTCTGATGATGAGGAGCAC 780
QY 867 CCTTGGAGTCTGACCTTCGCAACCG 894
Db 781 CCTTGGAGTCTGACCTTCGCAACCG 808

RESULT 15
CD514811 898 bp mRNA linear EST 06-JUN-2003
LOCUS      CD514811
DEFINITION AGENCOURT_14376196 NIH MGC 181 Homo sapiens cDNA clone
IMAGE:30356201 5', mRNA sequence.
CD514811
VERSION    CD514811.1 GI:31446529
KEYWORDS   EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE   1 (bases 1 to 898)
AUTHORS     NIH-MGC http://mgc.nci.nih.gov/.
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished
COMMENT     Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgabbs-remail.nih.gov
            Tissue Procurement: Dr. Michael Brownstein
            cDNA Library Preparation: Invitrogen Corp
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LNL at:
            http://image.lnl.gov
            Plate: NDMA477 row: 1 column: 10
            High quality sequence stop: 662.
            Location/Qualifiers
                1..898
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="IMAGE:30356201"
                /tissue_type="White Matter"
                /dev_stage="Unknown"
                /lab_host="DH10B-Ton A ( T1 and T5 phage resistances)"
                /clone_1lb="NIH MGC 181"
                /note="Vector: pCMV-SPORT6.1; Site 1: NotI; Site 2: EcoRV
                (destroyed): Library is oligo-dT primed and directionally
                cloned (EcoRV site is destroyed upon cloning). Average
                insert size 1.42 kb. Library was constructed by
                (Invitrogen). Note: this is a NIH_MGC Library."
FEATURES
source

```


BASE COUNT 194 a 231 c 310 g 158 t 5 others
ORIGIN

Query Match 46.4%; Score 753.4; DB 14; Length 898;
Best Local Similarity 99.9%; Pred. No. 6.1e-150;
Matches 754; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 23 CGCTGGGCGGGCCCGCGGGAGAGGGCGGGGAGCCCGATGCGCGGAG 82
Db 1 CGCTGGGCGGGCCCGCGGGAGAGGGCGGGGAGCCCGATGCGCGGAG 60
QY 83 CGAGGCGGAGATGCTTCGGCGGAGCGAAGACTGCGAGGCCCGCGGAGCCG 142
Db 61 CGAGGCGGAGATGCTTCGGCGGAGCGAAGACTGCGAGGCCCGCGGAGCCG 120
QY 143 ACCGTCGCGACCGAGCGGCTTCCTGATAGGGGTGAGCGCGGACTGCGAGCGGAAAT 202
Db 121 ACCGTCGCGACCGAGCGGCTTCCTGATAGGGGTGAGCGCGGACTGCGAGCGGAAAT 180
QY 203 CGACCGGTGTGAGGAAGATGATGAGTTGCTGGGAGCAGAAAGAGTGAACAGCGGAGC 262
Db 181 CGACCGGTGTGAGGAAGATGATGAGTTGCTGGGAGCAGAAAGAGTGAACAGCGGAGC 240
QY 263 GGAAGGTGTCTATCTGAGCCAGGACAGTTCTACAAAGTCTGACCGGACAGCAGAGG 322
Db 241 GGAAGGTGTCTATCTGAGCCAGGACAGTTCTACAAAGTCTGACCGGACAGCAGAGG 300
QY 323 CCAAGGCTTGAAAGGACAGTACATTTTGAACCATCCAGATGCTTTGATATGATTTGA 382
Db 301 CCAAGGCTTGAAAGGACAGTACATTTTGAACCATCCAGATGCTTTGATATGATTTGA 360
QY 383 TGACACGAGACTGTGAGAGACATGTGAGGGCAAAAGGTGAGGTGCCAGCTATGAT 442
Db 361 TGACACGAGACTGTGAGAGACATGTGAGGGCAAAAGGTGAGGTGCCAGCTATGAT 420
QY 443 TTGTGACACACTGAAAGTTACAGAGACACAGGTGTCTACCTGCGGACGTGTTCTGT 502
Db 421 TTGTGACACACTGAAAGTTACAGAGACACAGGTGTCTACCTGCGGACGTGTTCTGT 480
QY 503 TTGAGGGCATCTTGTGTCTTAACAGCCAGAGATCCGGAATGTTCCACTGGGCTCT 562
Db 481 TTGAGGGCATCTTGTGTCTTAACAGCCAGAGATCCGGAATGTTCCACTGGGCTCT 540
QY 563 TCGTGAACAAGCACTCCAGCTCAGGCTGTCTGAAAGTTCCCGGGAAGTGGCCGAG 622
Db 541 TCGTGAACAAGCACTCCAGCTCAGGCTGTCTGAAAGTTCCCGGGAAGTGGCCGAG 600
QY 623 GGAAGGACCTGAGCAGATCTGACGAGTACACCACTTCGTGAAGCCGGCTTCGAGG 682
Db 601 GGAAGGACCTGAGCAGATCTGACGAGTACACCACTTCGTGAAGCCGGCTTCGAGG 660
QY 683 AGTTCTGCTGCGCAGCAAAAGATGATGATGATCCACGAGAGTGAACATA 742
Db 661 AGTTCTGCTGCGCAGCAAAAGATGATGATGATCCACGAGAGTGAACATA 720
QY 743 TGGTTGCAATCAACCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 777
Db 721 TGGTTGCAATCAACCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 755
```

Search completed: November 25, 2003, 02:00:45
Job time : 3919.04 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 25, 2003, 02:01:02 ; Search time 3908.04 Seconds
(without alignments)
10099.814 Million cell updates/sec

Title: US-09-896-522-1
Perfect score: 1624
Sequence: 1 gtcggggtcgcctccgacatc.....ccaggtctgtgcgggacag 1624

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 22781392 seqs, 12152238056 residues

Word size : 15

Total number of hits satisfying chosen parameters: 61409

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :
EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estum:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estcom:*
17: em_gse_hum:*
18: em_gse_inv:*
19: em_gse_pin:*
20: em_gse_vrt:*
21: em_gse_fun:*
22: em_gse_mam:*
23: em_gse_mus:*
24: em_gse_pro:*
25: em_gse_rtd:*
26: em_gse_prg:*
27: em_gse_vrt1:*
28: gb_gse1:*
29: gb_gse2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	
1	732	45.1	898	14	CD514811 AGENCOURT	
2	727	44.8	1118	12	BM467984 AGENCOURT	
3	708	43.6	1044	12	BM917506 AGENCOURT	
C	4	701	43.2	941	13	BX372390 BX372390

5	698	43.0	916	12	BG826894	BG826894 602750978
6	688	42.4	730	12	BG770518	BG770518 602734236
7	677	41.7	725	12	BT256928	BT256928 602975534
8	659	40.6	869	10	BG491384	BG491384 602535670
9	653	40.2	883	13	B0838101	B0838101 AGENCOURT
10	650	40.0	1036	13	BQ072501	BQ072501 AGENCOURT
11	643	39.6	1201	13	BX394295	BX394295 BX394295
12	634	39.0	894	13	BX349263	BX349263 BX349263
13	630	38.8	1055	12	BW545603	BW545603 AGENCOURT
14	628	38.7	779	9	AU131406	AU131406 AU131406
15	623	38.4	753	9	AU122008	AU122008 AU122008
16	619	38.1	956	13	BQ179741	BQ179741 AGENCOURT
17	615	37.9	989	10	BQ390519	BQ390519 602416183
18	615	37.9	989	10	BQ390519	BQ390519 602416183
19	612	37.7	922	13	BQ935919	BQ935919 AGENCOURT
20	605	37.3	963	14	CD359372	CD359372 AGENCOURT
21	601	37.0	661	12	B1830878	B1830878 603081002
22	592	36.5	901	13	B1258532	B1258532 602972340
23	589	36.3	848	12	B1258532	B1258532 602972340
24	579	35.7	778	14	CD101693	CD101693 AGENCOURT
25	578	35.6	860	10	BG491358	BG491358 602535642
26	566	34.9	945	13	BQ677293	BQ677293 AGENCOURT
27	564	34.7	1039	12	BM803129	BM803129 AGENCOURT
28	562	34.6	922	10	BF664526	BF664526 602146091
29	561	34.5	948	13	BQ680950	BQ680950 AGENCOURT
30	558	34.4	827	12	BT261258	BT261258 602969110
31	558	34.4	959	12	BT770572	BT770572 603060368
32	556	34.2	796	9	AU141725	AU141725 AU141725
33	556	34.2	799	12	BG723578	BG723578 602694293
34	554	34.1	554	12	BM822058	BM822058 K-EST0091
35	554	34.1	928	10	BQ680187	BQ680187 AGENCOURT
36	542	33.4	998	10	BG492079	BG492079 602535670
37	537	33.1	797	12	BT116432	BT116432 602688854
38	535	32.9	998	13	BX400889	BX400889 BX400889
39	525	32.3	584	10	AW961647	AW961647 EST173720
40	522	32.1	1095	12	BM802939	BM802939 AGENCOURT
41	520	32.0	1114	13	BX433101	BX433101 BX433101
42	515	31.7	515	12	BM835278	BM835278 K-EST0110
43	514	31.7	514	12	BM827668	BM827668 K-EST0100
44	508	31.3	873	13	BQ877260	BQ877260 AGENCOURT
45	507	31.2	897	13	BQ681704	BQ681704 AGENCOURT

ALIGNMENTS

RESULT 1
LOCUS CD514811
DEFINITION AGENCOURT.14376196 NIH MGC.181 Homo sapiens cDNA clone
IMAGE:30356201 5', mRNA sequence.
CD514811
CD514811.1 GI:31446529
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgaabs-remail.nih.gov
Tissue Procurement: Dr. Michael Brownstein
cDNA Library Preparation: Invitrogen Corp
cDNA Library Arrayed By: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:


```

Db      310 TCCAGATGCTTTGATATGATTTGATGACAGAGACTCTGAGAAACATGTTGAGGGCAA 369
Qy      417 AAGGTGAGAGTTCGCCAGCTATGATTTTGTGACACACTCAAGTTACAGAGACACGGT 476
Db      370 AAGGTGAGAGTTCGCCAGCTATGATTTTGTGACACACTCAAGTTACAGAGACACGGT 429
Qy      477 GGTCTACCCCTGGGAGAGTGTCTGTTTGGAGGATCTTGTTGTTTACAGAGGAGAT 536
Db      430 GGTCTACCCCTGGGAGAGTGTCTGTTTGGAGGATCTTGTTGTTTACAGAGGAGAT 489
Qy      537 CCGGAGACATGTTCCACCTGCGCTCTTCTGTGACACCCGACTCCGACGTGCTCTCG 596
Db      490 CCGGAGACATGTTCCACCTGCGCTCTTCTGTGACACCCGACTCCGACGTGCTCTCG 549
Qy      597 AAGAGTTCTCCGAGACGTGCGCGGAGGAGGACCTTGAGACATTTTGAACGACATCAC 656
Db      550 AAGAGTTCTCCGAGACGTGCGCGGAGGAGGACCTTGAGACATTTTGAACGACATCAC 609
Qy      657 CACCTTCGTGAGCCGCTTCGAGAGATTCTGCTGCTCCGACAAAGAGATGCCGATGT 716
Db      610 CACCTTCGTGAGCCGCTTCGAGAGATTCTGCTGCTCCGACAAAGAGATGCCGATGT 669
Qy      717 GATCATCCACGAGAGAGTGGACATATGTTGCCATCACTGATCTGTCAGACATCCA 776
Db      670 GATCATCCACGAGAGTGGACATATGTTGCCATCACTGATCTGTCAGACATCCA 729
Qy      777 GGACATT 783
Db      730 GGACATT 736

```

```

RESULT 3
LOCUS    BM917506                      1044 bp      mRNA      linear      EST 12-MAY-2002
DEFINITION
AGENCOURT_6606568 NIH_MGC_106 Homo sapiens cDNA clone IMAGE:5484033
5', mRNA sequence.
BM917506
BM917506.1 GI:19367885
EST.
Homo sapiens (human)
SOURCE   Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1044)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Dr. Daniel McVicar, DBS/NCI
cDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.jnl.gov
Plate: LHCN2012 row: k column: 10
High quality sequence stop: 624.
Location/Qualifiers
1. 1044
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5484033"
/tissue_type="natural killer cells, cell line"
/lab_host="DH10B (phage-resistance)"
/clone_lib="NIH MGC 106"
/notes="Organ: blood; Vector: pORF7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGGACGAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and

```

FEATURES

source

```

1. 1044
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5484033"
/tissue_type="natural killer cells, cell line"
/lab_host="DH10B (phage-resistance)"
/clone_lib="NIH MGC 106"
/notes="Organ: blood; Vector: pORF7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGGACGAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and

```

```

BASE COUNT      224 a      306 c      318 g      196 t
ORIGIN
Superscript II RT (Life Technologies). Note: this is a
NIH MGC Library."

```

```

Query Match      43.6%; Score 708; DB 12; Length 1044;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 708; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      106 GAGAGCGAAAGTCTGCGAGAGCCCGCGGAGGCGGACCGTCCGACGAGCGGCCCTTC 165
Db      1   GAGAGCGAAAGTCTGCGAGAGCCCGCGGAGGCGGACCGTCCGACGAGCGGCCCTTC 60
Qy      166 CTGATGAGGGGTGAGCGCGCGGACCTGCGACCGGGGAAGTCCACCGTGTGTGAAGAAATCATG 225
Db      61   CTGATGAGGGGTGAGCGCGCGGACCTGCGACCGGGGAAGTCCACCGTGTGTGAAGAAATCATG 120
Qy      226 GAGTTGCTGGGACAGAAAGAGGTGGAACAGCGGACGAGAGGTGTCATCTGAGCCAG 285
Db      121 GAGTTGCTGGGACAGAAAGAGGTGGAACAGCGGACGAGAGGTGTCATCTGAGCCAG 180
Qy      286 GACAGGTTCTACAGAGTCTGACGCGACGACGAAAGGCCAGGCTTGAAGGACATAC 345
Db      181 GACAGGTTCTACAGAGTCTGACGCGACGACGAAAGGCCAGGCTTGAAGGACATAC 240
Qy      346 AATTTTGACATCCGATGCGCTTTGATATGATTTTGTGACACAGATCTGGAAGAACATC 405
Db      241 AATTTTGACATCCGATGCGCTTTGATATGATTTTGTGACACAGATCTGGAAGAACATC 300
Qy      406 GTGAGAGGCAAAACGATGAGGTGCGGACCTATGATTTTGTGACACACTCAAGGTTACCA 465
Db      301 GTGAGAGGCAAAACGATGAGGTGCGGACCTATGATTTTGTGACACACTCAAGGTTACCA 360
Qy      466 GAGACCAACGATGATCTTACCTGCGGACGTTCTGTTTGTGAGGACATCTTGTTCTTAC 525
Db      361 GAGACCAACGATGATCTTACCTGCGGACGTTCTGTTTGTGAGGACATCTTGTTCTTAC 420
Qy      526 AGCAGAGAAATCCGGGACATGTTTCCACTGCGGCTCTTGTGTGACACCGACCTCCGACGTC 585
Db      421 AGCAGAGAAATCCGGGACATGTTTCCACTGCGGCTCTTGTGTGACACCGACCTCCGACGTC 480
Qy      586 AGGCTGTCTGAAAGATTTCTCCGGGACGTCGCCCGAGGAGGAACTGTGAGCAGATTCTG 645
Db      481 AGGCTGTCTGAAAGATTTCTCCGGGACGTCGCCCGAGGAGGAACTGTGAGCAGATTCTG 540
Qy      646 AGCAGATACCAACCTTCTGTAAGCCGCTTGAAGAGTTCCTGCGGACCAAGAG 705
Db      541 AGCAGATACCAACCTTCTGTAAGCCGCTTGAAGAGTTCCTGCGGACCAAGAG 600
Qy      706 TATGCCGATGATATCATCCAGAGAGTGAACATATGTTGCCATCAACTGATCTG 765
Db      601 TATGCCGATGATATCATCCAGAGAGTGAACATATGTTGCCATCAACTGATCTG 660
Qy      766 CAGCAGATCAGAGACATCTGTAATGATGATCTGCAATGACACCGA 813
Db      661 CAGCAGATCAGAGACATCTGTAATGATGATCTGCAATGACACCGA 708

```

```

RESULT 4
LOCUS    BX372390                      941 bp      mRNA      linear      EST 08-MAY-2003
DEFINITION
BX372390 Homo sapiens B CELLS (RAMOS CELL LINE) COR 25-NORMALIZED
Homo sapiens cDNA clone CSDBL009YD15 3-PRIME, mRNA sequence.
BX372390
BX372390.1 GI:30456111
EST.
Homo sapiens (human)
SOURCE   Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 941)
Gruber, C., Jesse, J. and Polayes, D.
Full-length cDNA libraries and normalization

```

JOURNAL
COMMENT

Unpublished
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 4968.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0BA1046ZH05_CS04360_1&cluster=4968.r.
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Paradise Avenue Genoscope sequence ID : CS0BA1046ZH05_CS04360_1.
location/Qualifiers

FEATURES
source

1. 941
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DL009YD15"
/cell_line="RAMOS CELL LINE"
/cell_line="RAMOS CELL LINE"
/clone_id="Homo sapiens B CELLS (RAMOS CELL LINE) COT
25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

BASE COUNT
ORIGIN

Query Match 43.0%; Score 701; DB 13; Length 941;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 701; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

542 ACATGTTCCACCTCGCTCTTCTGTGAACACGACCTCCGAGCTGCTGTGAAGAG 601
Db ACATGTTCCACCTCGCTCTTCTGTGAACACGACCTCCGAGCTGCTGTGAAGAG 673
Qy TTCTCCGGAGCTGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 661
Db TTCTCCGGAGCTGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 672
Qy TCGTGAAGCCGCTTCGAGAGGTTCTGCTGCCGACAAAGATATGCCAGTGTGATCA 721
Db TCGTGAAGCCGCTTCGAGAGGTTCTGCTGCCGACAAAGATATGCCAGTGTGATCA 612
Qy TCCGACGAGAGTGACATATGTGTTGCCATCACTTGATCTGCACGACATCCAGACA 781
Db TCCGACGAGAGTGACATATGTGTTGCCATCACTTGATCTGCACGACATCCAGACA 552
Qy TTCTGAATGTAATCTGCAATATGCAACCGAGAGGTTCAATGGCGGAGCTACAGC 841
Db TTCTGAATGTAATCTGCAATATGCAACCGAGAGGTTCAATGGCGGAGCTACAGC 492
Qy GGACCTTTTCTGAGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 842
Db GGACCTTTTCTGAGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 433
Qy TGGAGTCCAGCAGACGACCCCACTGAGGGGCTGCCGAGCCCTCAGGGAGGTTCCGCC 373
Db TGGAGTCCAGCAGACGACCCCACTGAGGGGCTGCCGAGCCCTCAGGGAGGTTCCGCC 902
Qy TGGAGTCCAGCAGACGACCCCACTGAGGGGCTGCCGAGCCCTCAGGGAGGTTCCGCC 313
Db TGGAGTCCAGCAGACGACCCCACTGAGGGGCTGCCGAGCCCTCAGGGAGGTTCCGCC 962
Qy GGCATGTGTGTTCAGGAGCTGAGGCTGGGAGCGCCCAACCCACACCTGTTCTCTCG 1021
Db GGCATGTGTGTTCAGGAGCTGAGGCTGGGAGCGCCCAACCCACACCTGTTCTCTCG 312
Qy GCGCAGCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1081
Db GCGCAGCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 252
Qy GTGTCACTCAGACTCACTTCTGAGACCTGACAGGAGGTTCTGAGGTTTTCAGCCACT 1141
Db GTGTCACTCAGACTCACTTCTGAGACCTGACAGGAGGTTCTGAGGTTTTCAGCCACT 192

Qy 1142 TAGCTCGTTGCGGCTTAAAGATCCCTTAGGCACTGAGGAATGCCAGATGTCAG 1201

Db 132 TAGCTCGTTGCGGCTTAAAGATCCCTTAGGCACTGAGGAATGCCAGATGTCAG 73

Qy 1202 GAAGCTGGAGGCTTCTGTGAGGAATGTAGGACATTTAT 1242

Db 72 GAAGCTGGAGGCTTCTGTGAGGAATGTAGGACATTTAT 32

RESULT 5

602750978F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:4903819 5',

LOCUS

602750978F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:4903819 5',

DEFINITION

602750978F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:4903819 5',

ACCESSION

602750978F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:4903819 5',

VERSION

602750978F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:4903819 5',

KEYWORDS

602750978F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:4903819 5',

SOURCE

602750978F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:4903819 5',

ORGANISM

602750978F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:4903819 5',

REFERENCE

602750978F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:4903819 5',

AUTHORS

602750978F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:4903819 5',

TITLE

602750978F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:4903819 5',

JOURNAL

602750978F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:4903819 5',

COMMENT

602750978F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:4903819 5',

FEATURES

602750978F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:4903819 5',

source

602750978F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:4903819 5',

BASE COUNT

602750978F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:4903819 5',

ORIGIN

602750978F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:4903819 5',

Query Match

602750978F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:4903819 5',

Best Local Similarity

602750978F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:4903819 5',

Matches

602750978F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:4903819 5',

Conservative

602750978F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:4903819 5',

Mismatches

602750978F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:4903819 5',

Indels

602750978F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:4903819 5',

Accession	Version	Keywords	Source	Organism	Reference	Authors	Title	Journal	Comment
Db	212	GAGTTGCTGGGACAGAA	CGAGGTGGAA	CAGGGGCGACGGGAA	GGTGCATCTGAGCCAG	285			
Qy	226	GAGTTGCTGGGACAGAA	CGAGGTGGAA	CAGGGGCGACGGGAA	GGTGCATCTGAGCCAG	285			
Db	212	GAGTTGCTGGGACAGAA	CGAGGTGGAA	CAGGGGCGACGGGAA	GGTGCATCTGAGCCAG	271			
Qy	286	GACAGGTTCTTCAAGGTC	CTTGAACGGCAGAC	GAGAGGCGCAAGGCT	CTTGAAAGACAGTAC	345			
Db	272	GACAGGTTCTTCAAGGTC	CTTGAACGGCAGAC	GAGAGGCGCAAGGCT	CTTGAAAGACAGTAC	331			
Qy	346	AATTTTGACATCCAGATG	CGCTTTGGATATGAT	TTTGATGCAACAGACT	CTGAAAGACATC	405			
Db	332	AATTTTGACATCCAGATG	CGCTTTGGATATGAT	TTTGATGCAACAGACT	CTGAAAGACATC	391			
Qy	406	GTGGAAGGGCAAAA	CGGTGAGAGGTGCG	CCGACCTATGATTTT	TGTGACACACTCAAGTTACCA	465			
Db	392	GTGGAAGGGCAAAA	CGGTGAGAGGTGCG	CCGACCTATGATTTT	TGTGACACACTCAAGTTACCA	451			
Qy	466	GAGACCAAGGAGTCTAC	CGCTCGGACGATGGT	TCGTTGAAGGACATCT	GTGGTCTTAC	525			
Db	452	GAGACCAAGGAGTCTAC	CGCTCGGACGATGGT	TCGTTGAAGGACATCT	GTGGTCTTAC	511			
Qy	526	AGCCAGAGATCCGGGAC	ATGTTCCACCTGGGCT	CTTCGTTGACACACGAC	CTCCGACGTC	585			
Db	512	AGCCAGAGATCCGGGAC	ATGTTCCACCTGGGCT	CTTCGTTGACACACGAC	CTCCGACGTC	571			
Qy	586	AGGCTGTCTGAAGATT	CTCCGGGACGTGCG	CCGAGGAGGAGAC	CTTGAGACATTTCTG	645			
Db	572	AGGCTGTCTGAAGATT	CTCCGGGACGTGCG	CCGAGGAGGAGAC	CTTGAGACATTTCTG	631			
Qy	646	ACGCAGTACACCACTT	GTGTGAAAGCCGGC	CTTCCAGAGATTCTG	CTGCTGCCGACAAAGAG	705			
Db	632	ACGCAGTACACCACTT	GTGTGAAAGCCGGC	CTTCCAGAGATTCTG	CTGCTGCCGACAAAGAG	691			
Qy	706	TATGCCGATGTGATCAT	CCCAAGAGATGAGCA	TAAT 743					
Db	692	TATGCCGATGTGATCAT	CCCAAGAGATGAGCA	TAAT 729					
RESULT 6	BG770518								
LOCUS	BG770518	730 bp	mRNA	linear	EST 15-MAY-2001				
DEFINITION	6027342361 NIH_MGC_49	Homo sapiens	CDNA clone IMAGE:4859577 5',						
ACCESSION	BG770518		mRNA sequence.						
VERSION	BG770518.1	GI:14081171							
KEYWORDS	EST.								
SOURCE	Homo sapiens (human)								
ORGANISM	Homo sapiens								
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;								
AUTHORS	Mammalia; Euteheria; Primates; Catarrhini; Homnidae; Homo.								
TITLE	1 (bases 1 to 730)								
JOURNAL	NIH-MGC http://mhc.nci.nih.gov/ .								
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)								
	Unpublished								
	Contact: Robert Strausberg, Ph.D.								
	Email: csapba-remail.nih.gov								
	Tissue Procurement: ATCC/DCT/DTF								
	CDNA Library Preparation: Ling Hong/Rubin Laboratory								
	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)								
	DNA Sequencing by: Incyte Genomics, Inc.								
	Clone distribution: MGC clone distribution information can be								
	found through the I.M.A.G.E. Consortium/LNLN at:								
	http://image.lnl.gov								
	Plate: L1CMI715 row: h column: 10								
	High quality sequence stop:720.								

FEATURES
source

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4659577"
/tissue_type="melanotic melanoma, high MDR (cell line)"
/lab_host="DH10B (phage-resistant)"
/clone_id="NH_MGC_49"

```

```

/note="organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Size-selected >500bp for average insert size
1 kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
Library."
BASE COUNT      156 a      190 c      258 g      126 t
ORIGIN

```

Query Match	42.4%	Score 688	DB 12	Length 730
Best Local Similarity	100.0%	Pred. No. 0		
Matches 688	Conservative 0	Mismatches 0	Indels 0	Gaps 0
OY	46	GGGGAGAGGGCGGAGCGCGGGGACCCGATGTGGCGGGAGCGGAGCGGAGCGCGAGATGGCTTCGGCG	105	
Db	34	GGGGAGAGGGCGGGCGCGGGGGACCCGATGTGGCGGGAGCGGAGCGCGAGATGGCTTCGGCG	93	
OY	106	GGAGGCGAAGACTGTCGAGAGGCCCCGCGCCCGAGGGCCGACCGTCCGACCAAGCGGCCCTTTC	165	
Db	94	GGAGGCGAAGACTGTCGAGAGGCCCCGCGCCCGAGGGCCGACCGTCCGCAACAGCGGCCCTTTC	153	
OY	166	CTGATTAAGGGGGGAGACGGCGGGCACTGCGCAGCGGGGAAGTGCACCGGTGTGAGAAATCATG	225	
Db	154	CTGATTAAGGGGGGAGACGGCGGGCACTGCGCAGCGGGGAAGTGCACCGGTGTGAGAAATCATG	213	
OY	226	GAGTGTGCTGGGACAGAACGAGGTGGAAACAGCGCAGCGGAGAGTGTATCTTAGCGCAG	285	
Db	214	GAGTGTGCTGGGACAGAACGAGGTGGAAACAGCGCAGCGGAGAGTGTATCTTAGCGCAG	273	
OY	286	GACAGGTTCTTCAAGAGTCTCTGACGGCAGACAGACAGAGGCCAAGGCTTTGAAGAGCAAGTAC	345	
Db	274	GACAGGTTCTTCAAGAGTCTCTGACGGCAGACAGACAGAGGCCAAGGCTTTGAAGAGCAAGTAC	333	
OY	346	AATTTTACCACTCCAGATGCTTTGATTAATGATTTGATGCAACGAGACTGTGAAAGAAATC	405	
Db	334	AATTTTACCACTCCAGATGCTTTGATTAATGATTTGATGCAACGAGACTGTGAAAGAAATC	393	
OY	406	GTGAGAGGGCAAAAACGGTGGAGGTGCGACCTATGTTTGTGACACACTCAAGTTTACCA	465	
Db	394	GTGAGAGGGCAAAAACGGTGGAGGTGCGACCTATGTTTGTGACACACTCAAGTTTACCA	453	
OY	466	GAGACCAACGGTGTCTTACCCCTGCGGACGTGGTCTGTTTGAGGGCACTTGTGTCTTAC	525	
Db	454	GAGACCAACGGTGTCTTACCCCTGCGGACGTGGTCTGTTTGAGGGCACTTGTGTCTTAC	513	
OY	526	AGCCAGGAGNATCCGGGACATGTTCCACTGTGCTTCTTGTGTGACACCAACTCCGAGATC	585	
Db	514	AGCCAGGAGNATCCGGGACATGTTCCACTGTGCTTCTTGTGTGACACCAACTCCGAGATC	573	
OY	586	AGGCTGTCTCGAAGAGTTCTCCGGGACGTGCGCCGAGAGGAGGACCTGGAGCAATTTCTG	645	
Db	574	AGGCTGTCTCGAAGAGTTCTCCGGGACGTGCGCCGAGAGGAGGACCTGGAGCAATTTCTG	633	
OY	646	ACGCAAGTACCAACCTTTCGTGAAGCCGGCCTTCGAGGAGTTCTGCTCGGACCAAGAG	705	
Db	634	ACGCAAGTACCAACCTTTCGTGAAGCCGGCCTTCGAGGAGTTCTGCTCGGACCAAGAG	693	
OY	706	TATGCCGATGTGATCATCCCAAGAGAG	733	
Db	694	TATGCCGATGTGATCATCCCAAGAGAG	721	

RESULT	7
B1256928	
LOCUS	
DEFINITION	602975753f4F1 NIH_MGC_12 Homo sapiens CDNA clone IMAGE:5114816 5',
ACCESSION	B1256928
VERSION	B1256928
KEYWORDS	EST.
	GI:14811813

SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 725)
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC) JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: Incyte Genomics, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Plate: LLM11279 row: c column: 09
 High quality sequence stop: 707.
 Location/Qualifiers
 1..725
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5114816"
 /tissue_type="cervical carcinoma cell line"
 /lab_host="DH10B"
 /clone_lib="NIH_MGC_12"
 /note="Organ: cervix; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.4 kb. Library prepared by Life Technologies."
 BASE COUNT 175 a 181 c 224 g 145 t
 ORIGIN
 Query Match 41.7%; Score 677; DB 12; Length 725;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 677; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 197 GGAAGTGCACCGCTGTGAGAAATCATGAGCTTCTGGGACAGACAGAGTGGAAACAGC 256
 Db 1 GGAAGTGCACCGCTGTGAGAAATCATGAGCTTCTGGGACAGAAACAGAGTGGAAACAGC 60
 Oy 257 GGCAGCGGAAGTGGTCTCTGAGCCAGAGCAGAGTCTTCAAGAGTCTCTGACGCGCAGAGC 316
 Db 61 GGCAGCGGAAGTGGTCTCTGAGCCAGAGCAGAGTCTTCAAGAGTCTCTGACGCGCAGAGC 120
 Oy 317 AGAAGGCCAGAGCTTGAAGAGACAGTCAATTTTACCATCCAGATGCCCTTTGATATG 376
 Db 121 AGAAGGCCAGAGCTTGAAGAGACAGTCAATTTTACCATCCAGATGCCCTTTGATATG 180
 Oy 377 ATTTGATGCACAGACTCTGAAGAAACATCGTGAAGGGCAAAACGCTGAGAGTGCAGACT 436
 Db 181 ATTTGATGCACAGACTCTGAAGAAACATCGTGAAGGGCAAAACGCTGAGAGTGCAGACT 240
 Oy 437 ATGATTTTGTGACACTCAAGTTTACAGAGACACAGCGTGTCTACCTCTCGGACCTGG 496
 Db 241 ATGATTTTGTGACACTCAAGTTTACAGAGACACAGCGTGTCTACCTCTCGGACCTGG 300
 Oy 497 TTCTGTTTGAAGGCACTTGGTGTCTTACAGCCAGAGAGATCCGGAACATGTTCCACTGC 556
 Db 301 TTCTGTTTGAAGGCACTTGGTGTCTTACAGCCAGAGAGATCCGGAACATGTTCCACTGC 360
 Oy 557 GCCTCTTCTGAGACACGAGCTCCGAGCTCAGGCTGTCTCGAAGAGTTCTCCGGGACGTGC 616
 Db 361 GCCTCTTCTGAGACACGAGCTCCGAGCTCAGGCTGTCTCGAAGAGTTCTCCGGGACGTGC 420
 Oy 617 GCCAGAGGAGGACCTTGAGAGAGATTCTGACGACGACACCACTTGTGTAAGCCGGCCT 676
 Db 421 GCCAGAGGAGGACCTTGAGAGAGATTCTGACGACGACACCACTTGTGTAAGCCGGCCT 480
 Oy 677 TCGAGAGTTTTCCTGCGCAAAAGAAATATGCCATGTGATATCCACGAGAGAGTGG 736
 Db 736 TCGAGAGTTTTCCTGCGCAAAAGAAATATGCCATGTGATATCCACGAGAGAGTGG 736

Db 481 TCGAGAGTTTTCCTGCGCCGACAAAGAAATATGCCATGTGATATCCACGAGAGAGTGG 540
 Oy 737 ACATATGTTTGCCTTAACCTGATGCTGAGACATCCAGACATTTCTTAATGTTGACA 736
 Db 541 ACATATGTTTGCCTTAACCTGATGCTGAGACATCCAGACATTTCTTAATGTTGACA 600
 Oy 797 TCTGCAAAATGACACCGAGAGAGGTTCCAATGGGCGAGCTCAAGCGGACCTTTCTGAGC 856
 Db 601 TCTGCAAAATGACACCGAGAGAGGTTCCAATGGGCGAGCTCAAGCGGACCTTTCTGAGC 660
 Oy 857 CAGGGAGCACCTTGGG 873
 Db 661 CAGGGAGCACCTTGGG 677
 RESULT 8
 BG491384 869 bp mRNA linear EST 27-MAR-2001
 LOCUS 602535670F1 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:4684936 5',
 DEFINITION mRNA sequence.
 ACCESSION BG491384.1 GI:13452896
 VERSION BG491384.1
 KEYWORDS EST.
 SOURCE Homo sapiens
 ORGANISM Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 869)
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC) JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: DCTD/DTF
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Plate: LLM1493 row: k column: 17
 High quality sequence stop: 863.
 Location/Qualifiers
 1..869
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4684936"
 /tissue_type="amelanotic melanoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_41"
 /note="Organ: skin; Vector: pOT7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dt priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."
 BASE COUNT 197 a 223 c 291 g 156 t 2 others
 ORIGIN
 Query Match 40.6%; Score 659; DB 10; Length 869;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 659; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 46 GGGGAAGGGCGGGCGCCGAGGACCCGATGCGCGGAGCGGAGCCGAGATGCTTGGCG 105
 Db 4 GGGGAAGGGCGGGCGCCGAGGACCCGATGCGCGGAGCGGAGCCGAGATGCTTGGCG 63
 Oy 106 GAGGCGAAGACTGCGAGAGCCCCGCGCGGAGGCCGACCGCTCCGACCAAGCGGCCCTTC 165
 Db 64 GAGGCGAAGACTGCGAGAGCCCCGCGCGGAGGCCGACCGCTCCGACCAAGCGGCCCTTC 123

OY		166	CTGATAGGGGTGAGCGGGGGCATCTGCCACGGGGAAGTGACCCTGTGTGAAGAATCAATG	225
Dd		124	CTGATAGGGGTGAGCGGGGCACCTCCACGGGGAAGTGCACCCTGTGTGAAGAATCAATG	183
OY		226	GAGTTTGCTGGACAAGAACGAGGTGMAACAGCGGCACGCGAAGGTGTATCCTTGAGCCAG	285
Dd		184	GAGTTTGCTGGACAAGAACGAGGTGMAACAGCGGCACGCGAAGGTGTATCCTTGAGCCAG	243
OY		286	GACAGGTTCTACAGGTCTCTGACGGCACAGAACGAAAGGCCAAGGCTTTGAAAGACATAC	345
Dd		244	GACAGGTTCTACAGGTCTCTGACGGCACAGAACGAAAGGCCAAGGCTTTGAAAGACATAC	303
OY		346	AATTTTGACATCCAGATGCCCTTTGATAATGATTTTGATGACACAGACTCAAGGTTACCA	405
Dd		304	AATTTTGACATCCAGATGCCCTTTGATAATGATTTTGATGACACAGACTCTGMAAGAACATC	363
OY		406	GTGAGGGGCAAAAACGCTGGAGAGGTGGCGACCTATGATTTTGTGACACACTCAAGGTTACCA	465
Dd		364	GTGAGGGGCAAAAACGCTGGAGAGGTGGCGACCTATGATTTTGTGACACACTCAAGGTTACCA	423
OY		466	GAGACCAAGGTGTCTACCCCTGCGACGTGTGTTCTGTTTAGAGGCATCTTGTTGTTCTAC	525
Dd		424	GAGACCAAGGTGTCTACCCCTGCGACGTGTGTTCTGTTTAGAGGCATCTTGTTGTTCTAC	483
OY		526	AGCCAGAGATTCGGGACAATGTTCCACTGCGGCTTTCTGTGTGACACCGACTCCGACGTC	585
Dd		484	AGCCAGAGATTCGGGACAATGTTCCACTGCGGCTTTCTGTGTGACACCGACTCCGACGTC	543
OY		586	AGGCTGTCTCGAAGATTCTCCGGGACGTGCGCCGACGAGGAGGACCTTGAGACATTTCTG	645
Dd		544	AGGCTGTCTCGAAGATTCTCCGGGACGTGCGCCGACGAGGAGGACCTTGAGACATTTCTG	603
OY		646	ACGCAGTACACCACTCTGTGAAGCGGCGCTTCAGAGAGTTCTGCTGCGCCGACAAAGAA	704
Dd		604	ACGCAGTACACCACTCTGTGAAGCGGCGCTTCAGAGAGTTCTGCTGCGCCGACAAAGAA	662
RESULT 9		BUB38101		
LOCUS		BUB38101		
DEFINITION		BUB38101	883 bp	mRNA EST 16-OCT-2002
ACCESSION		BUB38101		
VERSION		BUB38101		
KEYWORDS		BUB38101.1	GI:24022496	
SOURCE		Homo sapiens	(human)	
ORGANISM		Homo sapiens		
TITLE		Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
REFERENCE		Mammalia; Euteria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS		1 (bases 1 to 883)		
JOURNAL		NIH-MGC http://mgc.nci.nih.gov/.		
COMMENT		National Institutes of Health, Mammalian Gene Collection (MGC)		
		Unpublished		
		Contact: Robert Strausberg, Ph.D.		
		Email: cga@db-rcmail.nih.gov		
		Tissue Procurement: DCD/DHP		
		cDNA Library Preparation: Rubin Laboratory		
		cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)		
		DNA Sequencing by: Agencourt Bioscience Corporation		
		Clone distribution: MGC clone distribution information can be		
		found through the I.M.A.G.E. Consortium/LNLN at:		
		http://image.lnl.gov		
		Plate: L16CM2326 row: C column: 04		
		High quality sequence stop:655.		

FEATURES

```

/organism="Homo sapiens"
/mol_type="mrna"
/db_xref="taxon:9606"
/clone="IMAGE:6088251"
/tissue_type="melanotic melanoma, cell line"
lab_host="DH10B (phage-resistant)"

```

/clone_11b="N1H MGC 112"
 /note="Organ: skin; Vector: pOTB7; Site 1: XhoI; Site 2:
 EcoRI; cDNA made by oligo-dT priming. Directionally cloned
 into EcoRI/XhoI sites using the following 5' adaptor:
 GGCCACGAG(C). Library constructed by Ling Hong in the
 laboratory of Gerald M. Rubin (University of California,
 Berkeley) using ZAP-cDNA Synthesis Kit (Stratagene) and
 Superscript II RT (Life Technologies). Note: this is a
 N1H MGC Library."
 BASE COUNT 188 a 240 c 272 g 183 t
 ORIGIN

Query Match	40.2%;	Score 653;	DB 13;	Length 883;
Best Local Similarity	99.7%;	Pred. No. 0;		
Matches 753; Conservative	0;	Mismatches 2;	Indels 0;	Gaps 0;

QY	457	AGGTTACAGAGACACAGGTGATCTACCTCGACGAGCTGTTCTGTTAGAGGATCTTG	516
Db	2	AGGTTACAGAGACACAGGTGATCTACCTCGACGAGCTGTTCTGTTAGAGGATCTTG	61
QY	517	GTGTTCTACAGCCAGAGATCCGGAGATGTTCCACCTGCGCTCTTCTGTGACACCGAC	576
Db	62	GTGTTCTACAGCCAGAGATCCGGAGATGTTCCACCTGCGCTCTTCTGTGACACCGAC	121
QY	577	TCCGACGTCAAGCTGTCTCCGAAGTCTCCGGACCTGCGCCGAGAGGAGACTTGAG	636
Db	122	TCCGACGTCAAGCTGTCTCCGAAGTCTCCGGAGCTGCGCCGAGAGGAGACTTGAG	181
QY	637	CAGATTCTGACGCGAGTACACCACTTGTGAAAGCGGCGCTTCGAGGAGTTTGCTGCGG	696
Db	182	CAGATTCTGACGCGAGTACACCACTTGTGAAAGCGGCGCTTCGAGGAGTTTGCTGCGG	241
QY	697	ACAAAGAAAGTATGCGCATGTGATCATCCACGAGAGTGAACAATATGTTGCCATCAAC	756
Db	242	ACAAAGAAAGTATGCGCATGTGATCATCCCGGAGAGATGGAACAATATGTTGCCATCAAC	301
QY	757	CTGATTCGTGACGACATTCACGAGACATTCGAAATGTGACACATCTGCAATATGSCACCGAGA	816
Db	302	CTGATTCGTGACGACATTCACGAGACATTCGAAATGTGACACATCTGCAATATGSCACCGAGA	361
QY	817	GGGTCCATATGGGCGGAGCTACAGCGGACCTTTTCTGAGCCAGGGGACCACTCTGGATG	876
Db	362	GGGTCCATATGGGCGGAGCTACAGCGGACCTTTTCTGAGCCAGGGGACCACTCTGGATG	421
QY	877	CTGACTCTTCGCAAAAGGTCACTATTGAGTCCAGCAGCAGACCCCACTGAGGGGCTGCC	936
Db	422	CTGACTCTTCGCAAAAGGTCACTATTGAGTCCAGCAGCAGACCCCACTGAGGGGCTGCC	481
QY	937	GAGGCTCAGGGCAGGTCTCCCGCGCCGCGCATGTGTTCAGGGATCTGAGCCTGGGGAGCGCC	996
Db	482	GAGGCTCAGGGCAGGTCTCCCGCGCGCGCATGTGTTCAGGGATCTGAGCCTGGGGAGCGCC	541
QY	997	CACCCACACCCACTGCTTCTCTTCGAGGCAACCCAGGGAGATGTTTACAGCGAGCCTTC	1056
Db	542	CACCCACACCCACTGCTTCTCTTCGAGGCAACCCAGGGAGATGTTTACAGCGAGCCTTC	601
QY	1057	CTCACTCAGAGAGTGGAAACTCAATGTGTCACTCAACTCAACTTCTGGGACACTGACA	1116
Db	602	CTCACTCAGAGAGTGGAAACTCAATGTGTCACTCAACTCAACTTCTGGGACACTGACA	661
QY	1117	GGCCTTCCTGAGGTTTTCAGGCACTTATAGGCTGTTGGGTTTAAAGATCCCTCTAGGTCA	1176
Db	662	GGCCTTCCTGAGGTTTTCAGGCACTTATAGGCTGTTGGGTTTAAAGATCCCTCTAGGTCA	721
QY	1177	CTGAGAAATGCCACAGAAATGTGACGAGAGCCTGGG	1211
Db	722	CTGAGAAATGCCACAGAAATGTGACGAGAGCCTGGG	756

RESULT 10

LOCUS B0072501

DEFINITION AGNECOURT 6838909 NIH MGC 122 Homo sapiens cDNA clone IMAGE:5761694

1036 bp

mRNA

linear

EST 02-APR-2002

5', mRNA sequence.
 ACCESSION BC072501
 VERSION BC072501.1 GI:19901547
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 1036)
 NIH-MGC <http://mgs.nci.nih.gov/>.
 TITLE NIH-MGC National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished
 AUTHORS Contact: Robert Strassberg, Ph.D.
 JOURNAL Email: cs9apbs-remail.nih.gov
 COMMENT Tissue Procurement: Life Technologies, Inc.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: <http://image.llnl.gov>
 Plate: LHAM12810 row: 1 column: 15
 High quality sequence stop: 644.
 Location/Qualifiers
 1..1036
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5761694"
 /lab_host="DH10B"
 /clone_lib="NIH_MGC_122"
 /note="Organ: pooled lung and spleen; Vector: pCMV-SportE; Site 1: NotI; Site 2: EcoRV (destroyed); RNA source anonymous pool of 24 week female lung; 16 week female spleen, and 20-22 week male spleens. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.4 kb. Insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 026. Note: this is a NIH MGC Library."
 BASE COUNT 221 a 285 c 346 g 184 t
 ORIGIN
 Query Match 40.0%; Score 650; DB 13; Length 1036;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 650; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 46 GGGGAAAGGGGCGGCGGGGAGCCCGATGCGGAGCGGAGCCGAGATGGCTTCGCG 105
 DB 16 GGGGAAAGGGGCGGCGGGGAGCCCGATGCGGAGCGGAGCCGAGATGGCTTCGCG 75
 QY 106 GGAGGCGAAGACTCGGAGAGCCCGCGCGGAGCCGCTCCGACACAGCGGCCCTTC 165
 DB 76 GGAGGCGAAGACTCGGAGAGCCCGCGCGGAGCCGCTCCGACACAGCGGCCCTTC 135
 QY 166 CTGATAGGGGAGCGGCGGCACTGCCAGCGGGAAGTGCACCGTGTGAGAAGATCATG 225
 DB 136 CTGATAGGGGAGCGGCGGCACTGCCAGCGGGAAGTGCACCGTGTGAGAAGATCATG 195
 QY 226 GAGTTGCTGGAGCAAGACGAGGTGAACAGCGGCAAGCGGAAGGTGTGATCTTGAGCCAG 285
 DB 196 GAGTTGCTGGAGCAAGACGAGGTGAACAGCGGCAAGCGGAAGGTGTGATCTTGAGCCAG 255
 QY 286 GACAGGTTCTCAAGGTCCTGACCGGAGAGAGCAAGGCGCAAGGCTTGAAGGACATAC 345
 DB 256 GACAGGTTCTCAAGGTCCTGACCGGAGAGCAAGGCGCAAGGCTTGAAGGACATAC 315
 QY 346 AATTTTGCATCCAGATGCTTTGATATGATTGATGACAGAGACTTGAAGAATC 405
 DB 316 AATTTTGCATCCAGATGCTTTGATATGATTGATGACAGAGACTTGAAGAATC 375
 QY 406 GTGAGGGCAAAACGTGAGAGTGCAGACTATGATTTGTGACACACTCAAGGTTACA 465

|||||
 DB 376 GTGAGGGCAAAACGTGAGAGTGCAGACTATGATTTGTGACACACTCAAGTTACCA 435
 QY 466 GAGACCAACGGTGTCTTCAACCTTGGGAGCGTGTCTGTGTAAGGCACTTGGTCTTAC 525
 DB 436 GAGACCAACGGTGTCTTCAACCTTGGGAGCGTGTCTGTGTAAGGCACTTGGTCTTAC 495
 QY 526 AGCCAGAGATCCGGGACATGTTCCACCTTGGGAGCGTGTCTGTGTAAGGCACTTGGGAG 585
 DB 496 AGCCAGAGATCCGGGACATGTTCCACCTTGGGAGCGTGTCTGTGTAAGGCACTTGGGAG 555
 QY 586 AGGCTGTCTGAAAGATTCTCCGGGAGCGTGTCTGTGTAAGGCACTTGGGAGCGTGTCTG 645
 DB 556 AGGCTGTCTGAAAGATTCTCCGGGAGCGTGTCTGTGTAAGGCACTTGGGAGCGTGTCTG 615
 QY 646 AGCAGATACCAACCTTGGGAGCGGCGCTTGGAGAGATTCTGCGCTCC 695
 DB 616 AGCAGATACCAACCTTGGGAGCGGCGCTTGGAGAGATTCTGCGCTCC 665
 RESULT 11
 BX394295
 LOCUS 1201 bp mRNA linear EST 13-MAY-2003
 DEFINITION BX394295 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
 ACCESSION CDNA clone CS0DC013YE13 5-PRIME, mRNA sequence.
 VERSION BX394295
 KEYWORDS BX394295.1 GI:30624219
 SOURCE EST.
 ORGANISM Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 1201)
 I.J.W.B., Gruber, C., Jesse, J. and Polyes, D.
 Full-length cDNA libraries and normalization
 Unpublished
 CONTACT: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: secref@genoscope.cns.fr Web: www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 4968.r For
 more information about this cluster, see
<http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0DC013AC070P1&cluster=4968.r>
 Feng Liang Email: liang@lifetech.com URL: <http://fulllength.invitrogen.com/> Invitrogen Corporation 1600
 Faraday Avenue Genoscope sequence ID: CS0DC013AC070P1.
 Location/Qualifiers
 1..1201
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DC013YE13"
 /tissue_type="NEUROBLASTOMA COT 25-NORMALIZED"
 /clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
 BASE COUNT 288 a 284 c 363 g 226 t 40 others
 ORIGIN
 Query Match 39.6%; Score 643; DB 13; Length 1201;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 693; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 139 GCCGACCGTCCGACCAAGGCGCCCTTCTGATAGGGGTAGCGGCGCACTGCGACGGG 198
 DB 133 GCCGACCGTCCGACCAAGGCGCCCTTCTGATAGGGGTAGCGGCGCACTGCGACGGG 191
 QY 199 AAGTCACCGTGTGTGAGAAATGATGAGTTGCTGGACAGAACGAGGTGAAACGCGG 258

```

Db      192 AAGTCGACCGTGTGTGAGAAATCATGAGTGTCTGGGACAGAAAGAGTGTGAAACACGCG 251
Qy      259 CAGCGGAGAGTGTGTCTGAGCCAGAGACAGGTTCTTCAAAAGTCTTGAACGGCAGACAG 318
Db      252 CAGGNAAGGTGTGTCTGAGCCAGGACAGGTTCTTCAAAAGTCTTGAACGGCAGACAG 311
Qy      319 AAGGCCAAGGCTTGAAGAAGACAGTAAATTTTGAACCATCCAGATGCTTGTGATATGAT 378
Db      312 AAGGCCAAGGCTTGAAGAAGACAGTAAATTTTGAACCATCCAGATGCTTGTGATATGAT 371
Qy      379 TTGATGACAGACCTCTGAAAGAACATGTGTGAGGCGAAACCGTGAAGTGTCCGACCTAT 438
Db      372 TTGATGACAGACCTCTGAAAGAACATGTGTGAGGCGAAACCGTGAAGTGTCCGACCTAT 431
Qy      439 GATTTTGTACACACTCAAGTTAACAGAACCCAGGTGTCTTACCTTGGGAGAGTGTGTT 498
Db      432 GATTTTGTACACACTCAAGTTAACAGAACCCAGGTGTCTTACCTTGGGAGAGTGTGTT 491
Qy      499 CTGTTTGAAGGACCTTGTGTGTCTTACAGCCAGGAGATCCGGGACATGTTCCACCTGCG 558
Db      492 CTGTTTGAAGGACCTTGTGTGTCTTACAGCCAGGAGATCCGGGACATGTTCCACCTGCG 551
Qy      559 CTCTCTGTGACACCGAATCCGACGTCAGGCTGTCTTGAAGAGTTCTCCGGGACGTGCGC 618
Db      552 CTCTCTGTGACACCGAATCCGACGTCAGGCTGTCTTGAAGAGTTCTCCGGGACGTGCGC 611
Qy      619 CGAGGAGGAGACCTGTGAGCAGATTTCTGACGACGACCACTTCTGTGAAGCCGGCTTC 678
Db      612 CGAGGAGGAGACCTGTGAGCAGATTTCTGACGACGACCACTTCTGTGAAGCCGGCTTC 671
Qy      679 GAGGAGTCTGACCTGCGGACAAAGAGATGTCGATGATCATCCACGAGGAGTGGAC 738
Db      672 GAGGAGTCTGACCTGCGGACAAAGAGATGTCGATGATCATCCACGAGGAGTGGAC 731
Qy      739 AATATGTTTCCATCAACTGATGTGTGACGACATCCAGACATTTGATGATG 798
Db      732 AATATGTTTCCATCAACTGATGTGTGACGACATCCAGACATTTGATGATG 791
Qy      799 TGCMAATGCAACCGAGAGGTTCCAAATGGGCGGA 832
Db      792 TGCMAATGCAACCGAGAGGTTCCAAATGGGCGGA 825

```

```

RESULT 12
BX349263      894 bp      mRNA      linear      EST 05-MAY-2003
LOCUS      BX349263 Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED
DEFINITION      Homo sapiens cDNA clone CS0DL009YD15 5-PRIME, mRNA sequence.
ACCESSION      BX349263
VERSION      BX349263.1 GI:30375361
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 (bases 1 to 894)
AUTHORS      Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE      Full-length cDNA libraries and normalization
JOURNAL      Unpublished
COMMENT      Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 4968.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0BAG057ZAI1.CS05457.1acluster-4968.r.
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/InvitrogenCorporation.1600
Fairaday Avenue Genoscope sequence ID : CS0BAG057ZAI1_CS05457_1.
Location/Qualifiers
1..894

```

FEATURES

source

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DL009YD15"
/cell_type="B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"
/cell_line="RAMOS CELL LINE"
/clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE) COT
25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT      192 a      242 c      253 g      206 t      1 others
ORIGIN
Query Match      39.0%; Score 634; DB 13; Length 894;
Best Local Similarity 99.9%; Pred. No. 8..9e-311;
Matches 684; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy      859 GGGGACACCCCTGGAGTGTGACCTCTGGCAAAAGGTCAATTGGAGTCCAGACAG 918
Db      35 GGGGACACCCCTGGAGTGTGACCTCTGGCAAAAGGTCAATTGGAGTCCAGACAG 94
Qy      919 CCCCACTGAAGGGCTCCGACCTCAAGGACAGGTCTCCGCGCATGTGTTCAGGG 978
Db      95 CCCCACTGAAGGGCTCCGACCTCAAGGACAGGTCTCCGCGCATGTGTTCAGGG 154
Qy      979 ACTGAGCTGGGAGCGCCCAACCCACCACTGCTTCTGTGGCGCACCCAGGGAGT 1038
Db      155 ACTGAGCTGGGAGCGCCCAACCCACCACTGCTTCTGTGGCGCACCCAGGGAGT 214
Qy      1039 GTTAGAGGAGGCGCTTCTCACTAGAGTGAATCAGATGTGCTCACTGACCTCAA 1098
Db      215 GTTAGAGGAGGCGCTTCTCACTAGAGTGAATCAGATGTGCTCACTGACCTCAA 274
Qy      1099 CTTCGTGGGACACTGACAGGCGTTCTGTGAGTTTTCAGCCACTTAACTGTTGCGGTTT 1158
Db      275 CTTCGTGGGACACTGACAGGCGTTCTGTGAGTTTTCAGCCACTTAACTGTTGCGGTTT 334
Qy      1159 AAGATCCCTCTTAGTGTCACTGAGAAATGCCACAAATGTGACGAAGCTGGAGGCTTC 1218
Db      335 AAGATCCCTCTTAGTGTCACTGAGAAATGCCACAAATGTGACGAAGCTGGAGGCTTC 394
Qy      1219 TGTGAGGAATGTGAGGACCACTTATTTGGGAAATGTGAGACAGCCTTACACTGTGCTG 1278
Db      395 TGTGAGGAATGTGAGGACCACTTATTTGGGAAATGTGAGACAGCCTTACACTGTGCTG 454
Qy      1279 CCTGATGTTTGTGTGACAGTGAACCCACAGTGGGAGAGAGTTTTCAGTCTGATCTGG 1338
Db      455 CCTGATGTTTGTGTGACAGTGAACCCACAGTGGGAGAGAGTTTTCAGTCTGATCTGG 514
Qy      1339 TTCTTACACACTCACACATTAATCTGAAAGTTTGTGAACAGTACTTTCTTTTAC 1398
Db      515 TTCTTACACACTCACACATTAATCTGAAAGTTTGTGAACAGTACTTTCTTTTAC 574
Qy      1399 AATGTCATGTCTGTCAATGTTTCTGTGTTCTGTTTATTAACAAGGCTGTGTGGCCT 1458
Db      575 AATGTCATGTCTGTCAATGTTTCTGTGTTCTGTTTATTAACAAGGCTGTGTGGCCT 634
Qy      1459 ACAACCTAATTCATGACCCAGTGTGTCAGTCCAGGCTGACCTTACAGATATGGG 1518
Db      635 ACAACCTAATTCATGACCCAGTGTGTCAGTGTGTCAGGCTGACCTTACAGATATGGG 694
Qy      1519 AGCCACTGAGGAGTGTTCCTCCCGC 1543
Db      695 AGCCACTGAGGAGTGTTCCTCCCGC 719

```

```

RESULT 13
BM545603      1055 bp      mRNA      linear      EST 20-FEB-2002
LOCUS      BM545603
DEFINITION      AGENCOURT 6500279 NIH_MGC_125 Homo sapiens cDNA clone IMAGE:5587864
5', mRNA sequence.

```

Accession	Version	Keywords	Source	Organism	Reference Authors	Title	Journal	Comment
BMS45603	EST.			Homo sapiens (human)	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	1 (bases 1 to 1055) NIH-MGC http://mgs.nci.nih.gov/ National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished	Contact: Robert Strausberg, Ph.D. Email: csapbs-r@mail.nih.gov Tissue Procurement: Invitrogen	CNA Library Arrayed by: The I.M.A.G.E. Consortium (ULNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ULNL at: http://image.llnl.gov Plate: LLM12358 row: a column: 17 High quality sequence stop: 579.
Features	Location/Qualifiers	Score	DB	Length	Indels	Gaps		
BASE COUNT	208 a 343 c 314 g 190 t							
ORIGIN								
Query Match	38.8%; Score 630; DB 12; Length 1055;							
Best Local Similarity	100.0%; Pred. No. 1e-308;							
Matches 630; Conservative	0; Mismatches 0; Indels 0; Gaps 0;							
Dn	46 GGAGGAAAGGGGGGGCGCCGGACCCGATGCCGGGGAGGCGAGAGCCGATGGCTTCGGCG	105						
Dn	1 GGGGAAGGGGGGGCGCGCGGGACCCGATGCCGGAGGCGAGAGCCGATGGCTTCGGCG	60						
Dn	106 GGAGGCGAAGACTGCGAGAAGCCCGCGCGCGAGCCGACCGTCCGCACCAGCGCCCTTC	165						
Dn	61 GGAGGCGAAGACTGCGAGAAGCCCGCGCGCGAGCGCGACCGTCCGCACCAGCGCCCTTC	120						
Dn	166 CTGATTAGGGGGGAGCGGGGCACTGCGCAGCGGGGAAGTGACCGGTGTGAAGAATCATG	225						
Dn	121 CTGATTAGGGGGGAGCGGGGCACTGCGCAGCGGGGAAGTGACCGGTGTGAAGAATCATG	180						
Dn	226 GAGTTGCTGGAGCAAGAACGAGGTGGAACAGCGCGCAGAGGTGTGATCTTGAGCCAG	285						
Dn	181 GAGTTGCTGGAGCAAGAACGAGGTGGAACAGCGCGCAGAGGTGTGATCTTGAGCCAG	240						
Dn	286 GACGAGTTCTACAAGGTCTCTGACCGCAGACAGCAAGGCGCAAGGCTTGAAGAAGACGTA	345						
Dn	241 GACGAGTTCTACAAGGTCTCTGACCGCAGACAGCAAGGCGCAAGGCTTGAAGAAGACGTA	300						
Dn	346 AATTTTGACATCCAGATGCTTTGATATATGATTTGATGACACAGGACTCTGAAGACATC	405						
Dn	301 AATTTTGACATCCAGATGCTTTGATATATGATTTGATGACACAGGACTCTGAAGACATC	360						
Dn	406 GTGAGAGGGCAAAACGGTAGAGTCCGACCTATGATTTTGTGACACACTCAAGGTACCA	465						
Dn	361 GTGAGAGGGCAAAACGGTAGAGTCCGACCTATGATTTTGTGACACACTCAAGGTACCA	420						

Qy	466	GAGACCAAGGAGGTCATCCCTCGCGGAGCGAGTGGTCTGTGAGGGGATCTTGATGTTCTAC	525
Db	421	GAGACCAAGGAGGTCATCCCTCGCGGAGCGAGTGGTCTGTGAGGGGATCTTGATGTTCTAC	480
Qy	526	AGCCAGAGATCCGGGACATGTTCCACCTGGCGCTCTTCTGTGACACCGACTCCGACGTC	585
Db	481	AGCCAGAGATCCGGGACATGTTCCACCTGGCGCTCTTCTGTGACACCGACTCCGACGTC	540
Qy	586	AGGCTGTCTGGAAGTTCCTCCGGGACGTGGCGCGGAGGAGGACCTTGAGACATTTGTG	645
Db	541	AGGCTGTCTGGAAGTTCCTCCGGGACGTGGCGCGGAGGAGGACCTTGAGACATTTGTG	600
Qy	646	ACGCAGTACACCACTTCTGTGAGCGGCGC	675
Db	601	ACGCAGTACACCACTTCTGTGAGCGGCGC	630
RESULT 14			
AUI31406			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
FEATURES			
source			
1..779			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/clone="NT2RP3002519"			
/cell_type="teratocarcinoma"			
/cell_line="NT2"			
/clone_id="NT2RP3"			
/note="Vector: pME18SFuj; mRNA from NT2 neuronal precursor cells after 2-weeks retinoic acid (RA) induction"			
BASE COUNT			
ORIGIN			
Query Match			
Match Local Similarity 99.9%; Pred. No. 9, 6e-308;			
Matches 678; Conservative 0; Mismatches 1; Indels 0; Gaps 0;			
Qy	71	GATCGCGGAGCGCGAGCGCGAGTGCCTTCGCGCGGAGCGAGACTGCGAGAGCCCG	130
Db	1	GATCGCGGAGCGCGAGCGCGAGTGCCTTCGCGCGGAGCGAGACTGCGAGAGCCCG	60
Qy	131	CGCCGAGGCCGACCGTCCGACCAACGCGGCCCTTCTGTATAGGGGTAGCGGCGGCACTG	190
Db	61	CGCCGAGGCCGACCGTCCGACCAACGCGGCCCTTCTGTATAGGGGTAGCGGCGGCACTG	120

```

QY 191 CCAGCGGGAAGTGCACCGTGTGTGAGAAGATCATGAGTGTGCTGGGACAGAACGAGGTGG 250
    |||
Db 121 CCAGCGGGAAGTGCACCGTGTGTGAGAAGATCATGAGTGTGCTGGGACAGAACGAGGTGG 180
QY 251 AACAGCGGACGCGGGAAGTGTGATCTCTGAGCCAGAGCAGAGTTCTACAGAGTCTGACGG 310
    |||
Db 181 AACAGCGGACGCGGGAAGTGTGATCTCTGAGCCAGAGCAGAGTTCTACAGAGTCTGACGG 240
QY 311 CAGAGCAGAAAGGCCAAGGCTTTGAAAGACAGTACAAATTTTGACATCCAGATGCTTTG 370
    |||
Db 241 CAGAGCAGAAAGGCCAAGGCTTTGAAAGACAGTACAAATTTTGACATCCAGATGCTTTG 300
QY 371 ATAATGATTTGATGACAGAGACTCTGAAGAACATCGGAGGCGCAAAACGGTGAAGGTGC 430
    |||
Db 301 ATAATGATTTGATGACAGAGACTCTGAAGAACATCGGAGGCGCAAAACGGTGAAGGTGC 360
QY 431 CGACCTATGATTTTGTGACACACTCAAGTTTACAGAGACAGAGTGTGCTACCTGCGG 490
    |||
Db 361 CGACCTATGATTTTGTGACACACTCAAGTTTACAGAGACAGAGTGTGCTACCTGCGG 420
QY 491 ACCTGCTTCTGTTTGAAGGATCTTGTGTTCTACAGCCAGAGATCCGGAATGTTCC 550
    |||
Db 421 ACCTGCTTCTGTTTGAAGGATCTTGTGTTCTACAGCCAGAGATCCGGAATGTTCC 480
QY 551 ACCTGCGCCTCTCTGAGGACACCGACTCCGACGTCAGGCTGTCTCGAAGATTCTCGGG 610
    |||
Db 481 ACCTGCGCCTCTCTGAGGACACCGACTCCGACGTCAGGCTGTCTCGAAGATTCTCGGG 540
QY 611 ACCTGCGCCTCTCTGAGGACACCGACTCTGAGCAGTACACCACTTGTGTAAGC 670
    |||
Db 541 ACCTGCGCCTCTCTGAGGACACCGACTCTGAGCAGTACACCACTTGTGTAAGC 600
QY 671 CGGCTTTCAGAGGATCTGCTGCTCCGACAAAGAAATGATGATGATATCCACGAG 730
    |||
Db 601 CGGCTTTCAGAGGATCTGCTGCTCCGACAAAGAAATGATGATGATATCCACGAG 660
QY 731 GAGTGGACAATATGTTGC 749
    |||
Db 661 GAGTGGACAATATGTTGC 679

```

```

RESULT 15
AUI22008 753 bp mRNA linear EST 01-AUG-2002
LOCUS AUI22008 MAMMA1 Homo sapiens cDNA clone MAMMA1001476 5', mRNA
DEFINITION sequence.
ACCESSION AUI22008
VERSION AUI22008.1 GI:10937243
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 753)
Ota.T., Nishikawa.T., Suzuki.Y., Ishii.S., Saito.K., Kawai.Y.,
Yamamoto.J., Wakamatsu.A., Nakamura.Y., Nagai.T., Sugano.S. and
Isogai.T.
TITLE HRI human cDNA project
JOURNAL Unpublished
COMMENT Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: genomics@hri.co.jp
HRI human cDNA project; 5'- & 3'-end one pass sequencing; Helix
Research Institute; cDNA library construction; Department of
Virology, Institute of Medical Science, University of Tokyo, and
Helix Research Institute.
Location/Qualifiers
1..753
/organism="Homo sapiens"

```

```

/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MAMMA1001476"
/tissue_type="mammary gland"
/clone_1b="MAMMA1"
/notes="Vector: PME18P1.3"
BASE COUNT 158 a 196 c 261 g 135 t 3 others
ORIGIN

```

```

Query Match 38.4%; Score 623; DB 9; Length 753;
Best Local Similarity 100.0%; Pred. No. 3.3e-305;
Matches 623; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 46 GGGGAAGGGGCGGCGCGGCGGACCGGATCGGGGAGCGGAGCCGAGATGCTTGGGG 105
    |||
Db 47 GGGGAAGGGGCGGCGGCGGACCGGATCGGGGAGCGGAGCCGAGATGCTTGGGG 106
QY 106 GGAAGCGAAGACTGCGAGAGCGCCGCGCGGAGCGGACCGTCCGACAGCGGCGCTTC 165
    |||
Db 107 GGAAGCGAAGACTGCGAGAGCGCCGCGCGGAGCGGACCGTCCGACAGCGGCGCTTC 166
QY 166 CTGATAGGGGTGAGCGGCGGCACTGCGACGCGGAACTGCACTGTGTGAGAAATCATG 225
    |||
Db 167 CTGATAGGGGTGAGCGGCGGCACTGCGACGCGGAACTGCACTGTGTGAGAAATCATG 226
QY 226 GAGTTGCTGGGACAGAACGAGGTGGAAACAGCGCGGAGGAAAGTGTGATCTGAGGCCAG 285
    |||
Db 227 GAGTTGCTGGGACAGAACGAGGTGGAAACAGCGCGGAGGAAAGTGTGATCTGAGGCCAG 286
QY 286 GACAGGTTCTAACAGGTCCTGACGCGGACAGAGCGCAAGAGGCGCTTGAAGGACAGTAC 345
    |||
Db 287 GACAGGTTCTAACAGGTCCTGACGCGGACAGAGCGCAAGAGGCGCTTGAAGGACAGTAC 346
QY 346 AATTTTGACCATCCAGATGCTTTGATTAATGATTTGATGATGATGATGATGATGATGATG 405
    |||
Db 347 AATTTTGACCATCCAGATGCTTTGATTAATGATTTGATGATGATGATGATGATGATGATG 406
QY 406 GTGAGAGGCAAAAGGTGAGGTGCGGACCTATGATTTTGTGACACACTCAAGATGATCA 465
    |||
Db 407 GTGAGAGGCAAAAGGTGAGGTGCGGACCTATGATTTTGTGACACACTCAAGATGATCA 466
QY 466 GAGACCAAGGTGATCTACACCTGCGGACGAGTGTCTGTTGAGGGGCACTTGGTGTCTAC 525
    |||
Db 467 GAGACCAAGGTGATCTACACCTGCGGACGAGTGTCTGTTGAGGGGCACTTGGTGTCTAC 526
QY 526 AGCCAGAGATCCGAGACATGTTCCACTGCGGCTTCTTGTGAGACACGACATCCGACGTC 585
    |||
Db 527 AGCCAGAGATCCGAGACATGTTCCACTGCGGCTTCTTGTGAGACACGACATCCGACGTC 586
QY 586 AGGCTGTCTCGAAGAGTTCTCCGGGACGTGCGCCGAGAGGAGGACCTGAGACAGATTTCTG 645
    |||
Db 587 AGGCTGTCTCGAAGAGTTCTCCGGGACGTGCGCCGAGAGGAGGACCTGAGACAGATTTCTG 646
QY 646 ACAGAGTACACCACTTGTGAA 668
    |||
Db 647 ACAGAGTACACCACTTGTGAA 669

```

```

Search completed: November 25, 2003, 06:51.11
Job time : 3921.54 secs

```

```

FEATURES
source

```

```

1..753
/organism="Homo sapiens"

```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus.n2p model

Run on: November 25, 2003, 07:39:21 ; Search time 16.8478 Seconds

(without alignments)
9066.011 Million cell updates/sec

Title: US-09-896-522-1

Perfect score: 536
Sequence: 1 gtcggggcgccgcgcaccc.....ccagctgtgtcgggcgccag 1624

Scoring table:

	OLIGO
Xgapop 60.0	Xgapext 60.0
Ygapop 60.0	Ygapext 60.0
Fgapop 6.0	Fgapext 7.0
Delop 6.0	Delext 7.0

Searched: 127863 seqs, 47026705 residues

Word size: 15

Total number of hits satisfying chosen parameters: 5

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: listing first 45 summaries

Command line parameters:

-MODE=frame.n2p.model -DEV=xlh
-Q/cg2.1/USPTO.spool/US09896522/runat_21112003_184144_3222/app_query.faeta_1.2830
-DB=SwissProt 41 -QMT=fastan -SUFFIX=n2p_oli.rsp -MINMATCH=0.1 -LOOPLU=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=quality -THR MIN=15 -ALIGN=15 -MODE=LOCAL
-OUTPM=pro -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09896522.@CGN_1.1.33@runat_21112003_184144_3222 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREDS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARYS

Result No.	Score	Query	Match length	ID	Description
1	277	51.7	277	1 UCK1_HUMAN	O9h47 homo sapien
2	45	8.4	277	1 UCK1_MOUSE	P5263 mus musculu
3	25	4.7	261	1 UCK1_HUMAN	O9bzx2 homo sapien
4	25	4.7	261	1 UCK2_MOUSE	O9pms9 mus musculu
5	15	2.8	260	1 UCK_DROME	O9vc99 drosophila

ALIGNMENTS

RESULT 1
ID UCK1_HUMAN STANDARD; PRT; 277 AA.
AC O9h47;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE Uridine-cytidine kinase 1 (EC 2.7.1.48) (UCK 1) (uridine monophosphokinase 1) (Cytidine monophosphokinase 1).

GN UCK1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
RX MEDLINE=21203813; PubMed=11306702;
RA Van Rompay A.R., Nord A., Linden K., Johansson M., Karlsson A.;
RT "Phosphorylation of uridine and cytidine nucleoside analogs by two
human uridine-cytidine kinases.";
RL Mol. Pharmacol. 59:1181-1186 (2001).
RN [2]
RP SEQUENCE FROM N.A.
RA Ho Y.S., Johnson R.K.;
RT "Human uridine kinase from prostate cancer cell line (LNCap).";
RL Submitted (Apr-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Mammary gland;
RA Itogai T., Ota T., Hayaishi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
RA Wagatsuma M., Hosoi T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Chida Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
RA Matsubae S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahara K., Masuno Y.,
RA Nimomiya K., Iwatsuki T.;
RT "NEO human cDNA sequencing project.";
RL Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Xin Y.R., Yu L., Zhao S.Y.;
RT "Cloning of a new human cDNA similar to Mus musculus uridine kinase
cDNA.";
RL Submitted (Feb-2002) to the EMBL/GenBank/DBJ databases.
CC - FUNCTION: Phosphorylates uridine and cytidine to uridine
monophosphate and cytidine monophosphate. Does not phosphorylate
deoxyribonucleosides or purine ribonucleosides. Can use ATP or GTP
as a phosphate donor. Can also phosphorylate cytidine and uridine
nucleoside analogs such as 6-azauridine, 5-fluorouridine, 4-
thiouridine, 5-bromouridine, N(4)-acetylcytidine, N(4)-
benzoylcytidine, 5-fluorocytidine, 2-thiocytidine, 5-
methylcytidine, and N(4)-anisoylcytidine.
CC - CATALYTIC ACTIVITY: ATP + uridine = ADP + UMP.
CC - CATALYTIC ACTIVITY: ATP + cytidine = ADP + CMP.
CC - PATHWAY: Pyrimidine salvage pathway.
CC - TISSUE SPECIFICITY: Ubiquitous.
CC - SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

Alignment Scores:

Pred. No.:	1e-272	Length:	277
Score:	277.00	Matches:	277
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	51.68%	Indels:	0
DB:	1	Gaps:	0

US-09-896-522-1 (1-1624) x UCK1_HUMAN (1-277)

```

QY 94 ATGGCTTCGGGGAGGCGGAAGACTGCGAGAGCCCGCGCGGAGCGGACCGTCCGCAC 153
    |||
Db 1 MetalaeseraIaGlyGlyGluAspCysGluSerProAlaPProGluAlaAspArgProHis 20
    |||
QY 154 CAGGGGCGCTTCCTGATPAGGGGTAGCGCGGCACTGCGACGGGGAAGTGCACCTGTGT 213
    |||
Db 21 GlnArgProPheLeuIleGlyValaSerGlyGlyThrAlaSerGlyLysSerThrValaCys 40
    |||
QY 214 GAGAAAGATCATGAGATTGCTGGAGCAGAACGAGGTGGAACGCGGACGGGAGAGTGATC 273
    |||
Db 41 GlnuylleuMetGluLeuLeuGlyGlnmsnGluValGluGlnArgGlnArgLysVala 60
    |||
QY 274 ATCCGAGCCAGCAGAGGTTCTACAGAGGTCCTGACGCGAGACAGAAAGCCAGGCTTGG 333
    |||
Db 61 IleuSerGlnAspArgPheThrLysValaLeuThrAlaGluGlnLysValaLysAlaLeu 80
    |||
QY 334 AAAAGACATGATCAATTTTGAACCATCCAGATCCCTTGTATATGATTTGATGACAGAGACT 393
    |||
Db 81 LysGlyGlnTyraAspPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThr 100
    |||
QY 394 CTGAGAAACATCGTGGAGGCGCAAAACGGTGGAGGTGCGGACCTTGTATTTGTGACACAC 453
    |||
Db 101 LeuLysAsnIleValGluGlyLysThrValaGluValaProThrTyraAspPheValaThrHis 120
    |||
QY 454 TCAAGGTTACAGAGACACCGTGAGTCTACCCCTGCGGACGTGTTCTGTTGAGGCGATC 513
    |||
Db 121 SerArgLeuProGluThrThrValaValaTyraProAlaAspValaValLeuPheGluGlyLe 140
    |||
QY 514 TTGGTGTCTACACCCAGGAGATCCGGGACATGTTCCACCTCGCCCTCTTGTGTGAGACAC 573
    |||
Db 141 LeuValPheTyraSerGlnGlnIleArgAspMetPheHisLeuArgLeuPheValaAspThr 160
    |||
QY 574 GACTCCGACGTGAGGCTGTCTCGAAGAGTTCTCCGGGACGTGCGCCGAGGAGGAGGACCTG 633
    |||
Db 161 AspSerAspValaGlyLeuSerArgArgValaLeuArgAspValaArgGlyArgAspLeu 180
    |||
QY 634 GAGCAGATTTGACGACAGTACACCACTTCGTGAAGCGGCTTCGAGAGATTGTGCTG 693
    |||
Db 181 GlnGlnIleLeuThrGlnTyraThrThrPheValaLysProAlaPheGluGluPheCysLeu 200
    |||
QY 694 CCGACAAAGAGTATGCCGATGTATCATCCACGAGAGAGTGGACATATGTGTCATC 753
    |||
Db 201 ProThrLysLysTyraAlaAspValaIleLeuProArgGlyValaAspAsnMetValaIle 220
    |||
QY 754 AACCTGATCGGACGACATCCAGACATTTGATGATGGTGACATCGCAAAATGGACACGGA 813
    |||
Db 221 AsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysGlySerThrHisArg 240
    |||
QY 814 GAGAGGTCGAATGGCGGAGGAGTCAACAGCGAACCTTTTCTGAGCGAGGAGACCACTGGG 873
    |||
Db 241 GlyGlySerAsnGlyArgSerTyraArgThrPheSerGlnPProGlyAspHisProGly 260
    |||
QY 874 ATGCTGACCTTGCGCAACGGTCACTTTGAGATGCCAGACAGACACCCAC 924
    |||
Db 261 MetLeuThrSerGlyArgSerHisLeuGlnSerSerSerArgProHis 277
    |||

```

RESULT 2

UCK1_MOUSE	STANDARD.	PRT.	277 AA.
ID_UCK1_MOUSE			
AC_P52623			
DT_01-OCT-1996	(Rel. 34, Created)		
DT_28-FEB-2003	(Rel. 41, Last sequence update)		
DT_15-SEP-2003	(Rel. 42, Last annotation update)		

DE Uridine-cytidine kinase 1 (EC 2.7.1.48) (UCK 1) (Uridine
 DE monophosphokinase 1) (Cytidine monophosphokinase 1).
 GN UCK1 OR UMPK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NX NCBI_TaxID=10090;
 RN [1]

RP SEQUENCE FROM N.A.
 RC TISSUE=Salivary gland;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Krausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Straussberg R.L., Collins P.S., Wagner L., Stenzen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buertow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Datchenko L., Marusita K., Farmer A.A., Rubin G.M., Hong L.,
 RA Scapleton M., Soares M.B., Bonaldo M.F., Casavani T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raba S.S., Loguellano N.A., Peters G.J., Abramson R.D., Mallary S.J.,
 RA Boek S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimmett J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
 RA Schmechel A., Schein U.E., Jones S.J.W., Maita W.A.,
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [2]

RP SEQUENCE OF 18-277 FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=97108719; PubMed=8951040;
 RA Ropp P.A., Traut T.W.,
 RT "Cloning and expression of a cDNA encoding uridine kinase from mouse
 RT brain.";
 RL Arch. Biochem. Biophys. 336:105-112 (1996).

CC -1- FUNCTION: Phosphorylates uridine and cytidine to uridine
 CC monophosphate and cytidine monophosphate. Does not phosphorylate
 CC deoxyribonucleosides or purine ribonucleosides. Can use ATP or GTP
 CC as a phosphate donor. Can also phosphorylate cytidine and uridine
 CC nucleoside analogs such as 6-azauridine, 5-fluorouridine, 4-
 CC thioridine, 5-bromouridine, N(4)-acetylcytidine, N(4)-
 CC benzoylcytidine, and N(4)-anisoylcytidine (By similarity).
 CC methylcytidine, and N(4)-anisoylcytidine (By similarity).
 CC -1- CATALYTIC ACTIVITY: ATP + uridine = ADP + UMP.
 CC -1- CATALYTIC ACTIVITY: ATP + cytidine = ADP + CMP.
 CC -1- PATHWAY: Pyrimidine salvage pathway.

CC -1- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: BC025146; AAH25146.1; -
 DR EMBL: L31783; AAB50568.1; -
 DR MGD: MGI:98804; Umpk.
 DR InterPro: IPR006083; PRK URK.
 DR InterPro: IPR000764; Uridine_kin.
 DR Pfam: PF00485; PRK.1.
 DR PRINTS: PR00988; URIDINKINASE.
 DR TIGRFAMs: TIGR00235; udk.1.
 KW Transferase; kinase; ATP-binding.
 FT NP BIND 30 ATP (POTENTIAL).
 SQ SEQUENCE 277 AA; 31068 MW; 3EBB3C4187FAEB4A CRC64;

Alignment Scores:


```

Pred. No.: 6,46e-37 Length: 277
Score: 45.00 Matches: 45
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 8.40% Indels: 0
Gaps: 0
DB: 1

US-09-896-522-1 (1-1624) x UCK1_MOUSE (1-277)

QY 391 ACCTGTAAGACATCGTGGAGGCGGAAACGCTGAGGTGCCACTTATGATTTTGACA 450
    |||
Db 100 ThrleuysanlleValIgluglyLysThrValgluValProThrTyraeapneValThr 119
    |||

QY 451 CACTCAAGTACAGAGACGACGAGTGTACCCCTGCGAGCGGTCTGTTGAGGCG 510
    |||
Db 120 HisSerArgleuProgluThrValValTyProAlaapValValLeupneglugly 139
    |||

QY 511 ATCTGTGTTCTTAC 525
    |||
Db 140 IleleuValpHeTy 144
    |||

RESULT 3
UCK2_HUMAN
ID UCK2_HUMAN STANDARD; PRT; 261 AA.
AC Q9BZX2; Q96KG5;
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last sequence update)
DE Uridine-cytidine kinase 2 (EC 2.7.1.48) (UCK 2) (Uridine
GN monophosphokinase 2) (Cytidine monophosphokinase 2).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A. AND CHARACTERIZATION.
MEDLINE=21203813; PubMed=11306702;
RA Van Rompay A.R., Norder A., Linden K., Johansson M., Karlsson A.;
RT "Phosphorylation of uridine and cytidine nucleoside analogs by two
RL human uridine-cytidine kinases."
RM Mol. Pharmacol. 59:1181-1186(2001).
[2]
SEQUENCE OF 15-261 FROM N.A.
RP TISSUE=Brain;
RC MEDLINE=21385121; PubMed=11494055;
RA Koizumi K., Shimamoto Y., Azuma A., Mataya Y., Matsuda A., Sasaki T.,
RA Fukushima M.;
RT "Cloning and expression of uridine/cytidine kinase cDNA from human
RT fibrosarcoma cells."
RL Int. J. Mol. Med. 8:273-278(2001).
CC -1- FUNCTION: Phosphorylates uridine and cytidine to uridine
CC monophosphate and cytidine monophosphate. Does not phosphorylate
CC deoxyribonucleosides or purine ribonucleosides. Can use ATP or GTP
CC as a phosphate donor. Can also phosphorylate cytidine and uridine
CC nucleoside analogs such as 6-azauridine, 5-fluorouridine, 4-
CC thioridine, 5-bromouridine, N(4)-acetylcytidine, N(4)-
CC benzoylcytidine, 5-fluorocytidine, 2-thiocytidine, 5-
CC methylcytidine, and N(4)-anisoylcytidine.
CC -1- CATALYTIC ACTIVITY: ATP + uridine = ADP + UMP.
CC -1- CATALYTIC ACTIVITY: ATP + cytidine = ADP + CMP.
CC -1- PATHWAY: Pyrimidine salvage pathway.
CC -1- TISSUE SPECIFICITY: Expressed in placenta.
CC -1- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

DR EMBL; AF236637; AAK14053.1; -
DR EMBL; AB062451; BAB56162.1; -
DR Genbank; HGNC:12562; UMPK.
DR InterPro; IPR006082; PRK.
DR InterPro; IPR006083; PRK_URK.
DR InterPro; IPR000764; Uridine_kin.
DR Pfam; PF00485; PRK_1.
DR PRINTS; PR00478; PHRIKINASE.
DR PRINTS; PR00988; URIDINKINASE.
DR TIGRFAMs; TIGR00235; udk; 1.
KW Transferase; Kinase; ATP-binding.
FT NP BIND 27 34 ATP (POTENTIAL).
SQ SEQUENCE 261 AA; 29299 MW; 71791346F091EBFD CRC64;

Alignment Scores:
Pred. No.: 1.39e-16 Length: 261
Score: 25.00 Matches: 25
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 4.66% Indels: 0
Gaps: 0
DB: 1

US-09-896-522-1 (1-1624) x UCK2_HUMAN (1-261)

QY 658 ACCTGTGTAAGCGGCGCTTCGAGAGTCTGCTGCCGACAAGATGCGATGNG 717
    |||
Db 187 ThrPheValysProAlaPheglugluPheCysleuProThrLysTyralaapVal 206
    |||

QY 718 ATCATCCACGAGGA 732
    |||
Db 207 IleleProArggly 211
    |||

RESULT 4
UCK2_MOUSE
ID UCK2_MOUSE STANDARD; PRT; 261 AA.
AC Q99PM9;
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last sequence update)
DE Uridine-cytidine kinase 2 (EC 2.7.1.48) (UCK 2) (Uridine
GN monophosphokinase 2) (Cytidine monophosphokinase 2).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
SEQUENCE FROM N.A.
RP MEDLINE=21203813; PubMed=11306702;
RA Van Rompay A.R., Norder A., Linden K., Johansson M., Karlsson A.;
RT "Phosphorylation of uridine and cytidine nucleoside analogs by two
RL human uridine-cytidine kinases."
RM Mol. Pharmacol. 59:1181-1186(2001).
CC -1- FUNCTION: Phosphorylates uridine and cytidine to uridine
CC monophosphate and cytidine monophosphate. Does not phosphorylate
CC deoxyribonucleosides or purine ribonucleosides. Can use ATP or GTP
CC as a phosphate donor. Can also phosphorylate cytidine and uridine
CC nucleoside analogs such as 6-azauridine, 5-fluorouridine, 4-
CC thioridine, 5-bromouridine, N(4)-acetylcytidine, N(4)-
CC benzoylcytidine, 5-fluorocytidine, 2-thiocytidine, 5-
CC methylcytidine, and N(4)-anisoylcytidine (by similarity).
CC -1- CATALYTIC ACTIVITY: ATP + uridine = ADP + UMP.
CC -1- CATALYTIC ACTIVITY: ATP + cytidine = ADP + CMP.
CC -1- PATHWAY: Pyrimidine salvage pathway.
CC -1- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

CC EMBL; AF236636; AAK14052.1; -
DR MGD; MG1:1931744; UCK2.
DR InterPro; IPR006082; PRK.
DR InterPro; IPR006083; PRK URK.
DR InterPro; IPR000764; Uridine_kin.
DR Pfam; PF00485; PRK; 1.
DR PRINTS; PR00478; PHRIBKINASE.
DR PRINTS; PR00988; URIDINKINASE.
DR TIGRFAMs; TIGR00235; uck; 1.
KM Transferase; Kinase; ATP-binding.
FT NP BIND 27 34 ATP (POTENTIAL).
SQ SEQUENCE 261 AA; 29404 MW; 780AA3CF5CA8153 CRC64;

Alignment Scores:
Pred. No.: 1,39e-16 Length: 261
Score: 25.00 Matches: 25
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 4.66% Indels: 0
DB: Gaps: 0

US-09-896-522-1 (1-1624) x UCK2_MOUSE (1-261)

Cy 658 ACCTTGCGAAGCGCGCTTGAGAGAGTTCGCTGCGGCAAGAGTATGCCAGATGTG 717
Db 187 ThreVallylsProAlaPheGluGluPheCysLeuProTrpLysIlyTAlaAspVal 206

Cy 718 ATCATCCACGAGGA 732
Db 207 lIeIleProArGly 211

RESULT 5
ID UCK_DROME STANDARD; PRF; 260 AA.
AC Q9VC99;
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Probable uridine-cytidine kinase (EC 2.7.1.48) (UCK) (Uridine
DE monophosphokinase) (Cytidine monophosphokinase).
GN CG6364.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoekins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abil J.F., Agbayant A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foeller C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasner K.,
RA Glodok A., Gong P., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laeko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,

```

```

RA Liu X., Matrei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milenina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacle J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Slier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
CC -1- CATALYTIC ACTIVITY: ATP + uridine = ADP + UMP.
CC -1- CATALYTIC ACTIVITY: ATP + cytidine = ADP + CMP.
CC -1- PATHWAY: Pyrimidine salvage pathway.
CC -1- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE003747; AAF56274.1; -
DR FlyBase; FBgn0039179; CG6364.
DR InterPro; IPR006083; PRK URK.
DR InterPro; IPR000764; Uridine_kin.
DR Pfam; PF00485; PRK; 1.
DR PRINTS; PR00988; URIDINKINASE.
KM Hypothetical protein; Transferase; Kinase; ATP-binding.
FT NP BIND 34 41 ATP (POTENTIAL).
SQ SEQUENCE 260 AA; 29213 MW; 2EE57423704925E6 CRC64;

Alignment Scores:
Pred. No.: 2.03e-06 Length: 260
Score: 15.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.80% Indels: 0
DB: Gaps: 0

US-09-896-522-1 (1-1624) x UCK_DROME (1-260)

Cy 481 TACCTCGGAGCGAGTTCGTTGAGGCGATCTGTGTCTAC 525
Db 135 TyPProAlaAspValValLeuPheGluGlyIleLeuValPheIlyr 149

Search completed: November 25, 2003, 07:54:01
Job time : 19.8478 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: November 25, 2003, 07:40:49 ; Search time 76.9715 Seconds

(without alignments)
10889.155 Million cell updates/sec

Title: US-09-896-522-1
Perfect score: 536
Sequence: 1 gtcggggctgcgcctccgaccc.....ccagcttctgaggggcag 1624

Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 830525 seqs, 258052604 residues

Word size: 15

Total number of hits satisfying chosen parameters: 6

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODE=frame+n2p.model -DEV=xlh
-Q=/cgm2_1/USPTO.spool/US09896522/runcat_21112003_184144_3228/app_query.fasta_1.2830
-DB=SPREMBL_23 -QFMT=fastan -SUFFIX=n2p_ol1.rspt -MISMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bites -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOCALLIN=200 -THR SCORE=quality -THR MIN=15 -ALIGN=15 -MODE=LOCL
-OUTFMT=pro -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09896522.QCGN_1_1_182.Qrunat_21112003_184144_3228 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

SPREMBL_23: *
1: sp_archaea: *
2: sp_bacteria: *
3: sp_fungi: *
4: sp_human: *
5: sp_invertebrate: *
6: sp_mammal: *
7: sp_mhc: *
8: sp_organelle: *
9: sp_phage: *
10: sp_plant: *
11: sp_rodent: *
12: sp_virus: *
13: sp_vertebrate: *
14: sp_unclassified: *
15: sp_virus: *
16: sp_bacteriaph: *
17: sp_archaeap: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	169	31.5	201	4 Q96BJ0	Q96bj0 homo sapien

2	25	4.7	105	11 Q9QY8	Q9QY8 ratcus norv
3	25	4.7	111	4 Q9BU42	Q9BU42 homo sapien
4	25	4.7	261	11 Q8C476	Q8C476 mus musculu
5	15	2.8	111	4 Q92528	Q92528 homo sapien
6	15	2.8	260	5 Q8WRJ1	Q8WRJ1 drosophila

ALIGNMENTS

RESULT 1

Q96BJ0 PRELIMINARY; PRT; 201 AA.
AC Q96BJ0;
DT 01-DEC-2001 (T-EMBLrel. 19, Created)
DT 01-DEC-2001 (T-EMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE Similar to uridine-cytidine kinase 1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Colon;
RA Strausberg R.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC015547; AAH15547.1; -.
DR InterPro; IPR006083; PRK_URK.
DR RefSeq; PR00485; PRK_1.
DR Pfam; PF00485; PRK_1.
DR PRINTS; PR00988; URIDINKINASE.
KW kinase.
SQ SEQUENCE 201 AA; 22790 MW; 0B5F2F00F7B5363 CRC64;

Alignment Scores:

Pred. No.: 5.61e-165 Length: 201
Score: 169.00 Matches: 169
Percent Similarity: 100.00% Conservative: 0
Beet Local Similarity: 100.00% Mismatches: 0
Query Match: 31.53% Indels: 0
DB: Gaps: 0

US-09-896-522-1 (1-1624) x Q96BJ0 (1-201)

QY	94	ATGGCTTCGGCGGAGCGCAAGACTCGGACCCCGCGGAGCCGACCGTCCGCAC	153
DB	1	MetAlaSerAlaGlyGlyGluAspCysGluSerProAlaProGluValAspArgProHis	20
QY	154	CAGCGGCTTCCTGATAGGGGTGAGCGGCGACCTGCGGAGGGAAGTCGACCGTGTG	213
DB	21	GlnArgProPheLeuIleGlyValSerGlyThrAlaSerGlyLysSerThrValCys	40
QY	214	GAGAGATCATGAGTTCGTGGAGAGAAAGAGTGGAACAGCGGACCGGAGAGTGTC	273
DB	41	GluValIleMetGluLeuLeuGlyGlnaengluValGluGlnArgGlnArgValVal	60
QY	274	ATCTGAGCCGAGACAGGTTCTTCAAGGCTCTGACGCGCAGACAGAGAGCCAAAGCTT	333
DB	61	IleLeuSerGlnAspArgPheThrLysValLeuThrAlaGluGlnLysAlaValLeu	80
QY	334	AAAGGACATGACATTTTGAACATCCAGATGCTTGAATGAATTTGATGACAGGACT	393
DB	81	LysGlyGlnThrAspPheAspHisProAspAlaPheAspAspLeuMetHisArgThr	100
QY	394	CTGAAGACATCGTGGAGGCAAAACGTTGAGTGCCGACCATGATTTTGTGACACAC	453
DB	101	LeuLysAsnIleValGluGlyLysThrValGluValProThrTyrAspPheValThrHis	120
QY	454	TCAAGTTTCCAGAGACACCGGTGCTTACCTCTGCGGACGTCGTCTGTTTGAAGGCATC	513
DB	121	SerArgLeuProGluThrThrValValTyrProAlaAspValValLeuPheGluGlyIle	140
QY	514	TTGGTGTCTTACAGCAGAGAGATCCGGGACATGTTCCACCTGCGGCTCTGTGACACAC	573

```
DB 141 Leuvalthyserginsulinlearghspethethestaurgleuphevalaspthr 160
QY 574 GACTCCGACGTCAGCTGTCTCGAGA 600
DB 161 Asperasapvalargleuserargarg 169

RESULT 2
Q9QY68 PRELIMINARY; PRT; 105 AA.
ID Q9QY68
AC Q9QY68
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Uridine Kinase (Fragment).
GN UK.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SD; TISSUE=Spinal cord;
RX MEDLINE=20050059; PubMed=10581173;
RA Yuh I., Yaoi T., Matsunabe S., Okajima S., Hirasawa Y., Fushiki S.;
RT "Up-regulated uridine kinase gene identified by RUCS in the ventral
RT horn after crush injury to rat sciatic nerves.";
RL Biochem. Biophys. Res. Commun. 266:104-109(1999).
DR EMBL; AB030700; BAA83085.1; -.
DR InterPro; IPR000764; Uridine_kin.
DR Pfam; PF00485; PRK.1.
DR PRINTS; PR00988; URIDINKINASE.
DR KINASE.
FT NON_TMR 1 1
FT NON_TMR 105 105
SQ SEQUENCE 105 AA; 11967 MW; 9DB43C15E07EB029 CRC64;

Alignment Scores:
Pred. No.: 3.66e-16 Length: 105
Score: 25.00 Matches: 25
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 4.66% Indels: 0
DB: 11 Gaps: 0

US-09-896-522-1 (1-1624) x Q9QY68 (1-105)
QY 658 ACCTTCGTGAAGCGGCTTCGAGAGATTCTGCTGCCGACAAGAAGTATGCCGATGTG 717
DB 36 Thirhevalysproalaphneglunphecysleuprotthylslytralaspval 55
QY 718 ATCATCCACGAGAGA 732
DB 56 llelleproarggly 60

RESULT 3
Q9BU42 PRELIMINARY; PRT; 111 AA.
ID Q9BU42
AC Q9BU42
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Similar to uridine monophosphate kinase.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RA Strauberg R.;
DT Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
```

```
DR EMBL; BC002906; AAH02906.1; -.
DR InterPro; IPR006083; PRK_URK.
DR InterPro; IPR000764; Uridine_kin.
DR Pfam; PF00485; PRK.1.
DR PRINTS; PR00988; URIDINKINASE.
DR KINASE.
SQ SEQUENCE 111 AA; 12587 MW; E6688B1B86F432A9 CRC64;

Alignment Scores:
Pred. No.: 3.64e-16 Length: 111
Score: 25.00 Matches: 25
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 4.66% Indels: 0
DB: 4 Gaps: 0

US-09-896-522-1 (1-1624) x Q9BU42 (1-111)
QY 658 ACCTTCGTGAAGCGGCTTCGAGAGATTCTGCTGCCGACAAGAAGTATGCCGATGTG 717
DB 37 Thirhevalysproalaphneglunphecysleuprotthylslytralaspval 56
QY 718 ATCATCCACGAGAGA 732
DB 57 llelleproarggly 61

RESULT 4
Q8C476 PRELIMINARY; PRT; 261 AA.
ID Q8C476
AC Q8C476
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Uridine-cytidine Kinase 2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium.
RA The Riken Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK082837; BAC38646.1; -.
SQ SEQUENCE 261 AA; 29438 MW; 780AA3C298AA8153 CRC64;

Alignment Scores:
Pred. No.: 3.27e-16 Length: 261
Score: 25.00 Matches: 25
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 4.66% Indels: 0
DB: 11 Gaps: 0

US-09-896-522-1 (1-1624) x Q8C476 (1-261)
QY 658 ACCTTCGTGAAGCGGCTTCGAGAGATTCTGCTGCCGACAAGAAGTATGCCGATGTG 717
DB 187 Thirhevalysproalaphneglunphecysleuprotthylslytralaspval 206
QY 718 ATCATCCACGAGAGA 732
DB 207 llelleproarggly 211

RESULT 5
Q92528 PRELIMINARY; PRT; 111 AA.
ID Q92528
AC Q92528
DT 01-FEB-1997 (TREMBlrel. 02, Created)
```

DT 01-FEB-1997 (TReMBLrel. 02, last sequence update)
 DT 01-MAR-2003 (TReMBLrel. 23, last annotation update)
 DE 5'-terminal region of UTK.
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Testis;
 RX MEDLINE=96411689; PubMed=8812458;
 RA Ozaki K., Kuroki T., Hayashi S., Nakamura Y.,
 RT "Isolation of three testis-specific genes (TSA303, TSA806, TSA903) by
 RT a differential mRNA display method.";
 RL Genomics 36:316-319 (1996).
 DR EMBL: D78335; BAA1349.1;
 DR InterPro: IPR006083; PRK_URK.
 DR InterPro: IPR000764; Uridine_kin.
 DR Pfam: PF00485; PRK; 1.
 DR PRINTS: PR00988; URIDINKINASE.
 SQ SEQUENCE 111 AA; 12617 MW; 81F6E4019CF428AB CRC64;

Alignment Scores:
 Pred. No.: 7.81e-06 Length: 111
 Score: 15.00 Matches: 15
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 2.80% Indels: 0
 DB: 4 Gaps: 0

US-09-896-522-1 (1-1624) x Q92528 (1-111)

QY 658 ACCTGTGTAAGCGGCTTCGAGAGCTTGTGCTGCCGCAAG 702
 DB 37 ThPheValIysProAlaPheGluGluPheCysLeuProThrIys 51

RESULT 6

Q8MRJ1 PRELIMINARY; PRT; 260 AA.
 AC Q8MRJ1;
 DT 01-OCT-2002 (TReMBLrel. 22, Created)
 DT 01-OCT-2002 (TReMBLrel. 22, last sequence update)
 DT 01-MAR-2003 (TReMBLrel. 23, last annotation update)
 DE LD13909p.
 GN CG6364.
 OS Drosophila melanogaster (fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkely;
 RA Stapleton M., Broksstein P., Hong L., Agbayan A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Fise E.,
 RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,
 RA Miranda A., Mungall C.J., Nunoo J., Paclab J., Paragas V., Park S.,
 RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.B., Rubin G.M.,
 RA Celisner S.;
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
 RL EMBL: AY19583; AA550237.1; -;
 DR FlyBase: FBgn0039179; CG6364.
 DR InterPro: IPR006083; PRK_URK.
 DR InterPro: IPR000764; Uridine_kin.
 DR Pfam: PF00485; PRK; 1.
 DR PRINTS: PR00988; URIDINKINASE.
 SQ SEQUENCE 260 AA; 29223 MW; 03CAA6DCA04A16S5 CRC64;

Alignment Scores:
 Pred. No.: 7.03e-06 Length: 260
 Score: 15.00 Matches: 15
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0

Query Match: 2.80% Indels: 0
 DB: 5 Gaps: 0

US-09-896-522-1 (1-1624) x Q8MRJ1 (1-260)

QY 481 TACCCTGCGAGCGTGTGTTTGAGGCGCATCTGTGTTCTAC 525
 DB 135 TyrProAlaAspValValLeuPheGluGluIleLeuValPheTyr 149

Search completed: November 25, 2003, 07:58:08
 Job time: 79.9715 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: November 25, 2003, 07:43:14 ; Search time 30.7225 Seconds
(without alignments)
10166.994 Million cell updates/sec

Title: US-09-896-522-1
Perfect score: 536
Sequence: 1 gtgggggtgcgcgcgcaccc.....ccaggtgtgtgggggacag 1624

Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 283308 seqs, 96168682 residues

Word size: 15

Total number of hits satisfying chosen parameters: 0

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: listing first 45 summaries

Command line parameters:

-MODE=frame+n2p.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US0896522/runat_21112003_184145_3234/app_query.fasta_1.2830
-DB=PIR_76 -QFMT=fastan -SUFFIX=n2p_ol1.rpr -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi -LIST=45
-DOCCALIGN=200 -THR SCORE=quality -THR MIN=15 -ALIGN=15 -MODE=LOCAL -OUTFMT=pro
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US0896522.@CGN_1_1_68.@runat_21112003_184145_3234 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : PIR_76:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB ID	Description

No matches found

Search completed: November 25, 2003, 07:59:52
Job time : 30.7225 secs

This Page Blank (uspto)

..... (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: November 25, 2003, 07:32:19 ; Search time 55.4988 Seconds

(without alignments)
9289.284 Million cell updates/sec

Title: US-09-896-522-1

Perfect score: 536

Sequence: 1 gtcgggtcgcctccgcagctc.....ccagctctgtcggggcag 1624

Scoring table:

OLIGO

Xgapop 60.0 , Xgapext 60.0

Ygapop 60.0 , Ygapext 60.0

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 1107863 seqs, 158726573 residues

Word size: 15

Total number of hits satisfying chosen parameters: 18

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODE=frame+n2p.model -DEV=xlh

-Q=/cqn2_1/USPTO.spool/US09896522/ragat_21112003_184143_3216/app_query.fasta_1.2830

-DB=A.Geneseq_19jun03 -QFMT=fastan -SUFFIX=n2p_011.rag -MINMATCH=0.1 -LOOPCL=0

-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi

-LIST=45 -DOCALLIGN=200 -THR SCORE=quality -THR MIN=15 -ALIGN=15 -MODE=LOCAL

-OUTFMT=pro -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000

-USER=US09896522.QCEN_1_1_112_@runat_21112003_184143_3216 -NCPU=6 -ICPU=3

-NO_MMAP -LARGEQUERY -MEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG

-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6

-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

A.Geneseq_19jun03.*

1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*

2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*

3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*

4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*

5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*

6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*

7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*

8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*

9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*

10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*

11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*

12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*

13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*

14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*

15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*

16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*

17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*

18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*

19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*

20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*

21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*

22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*

23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	296	55.2	296	22	AA041288
2	277	51.7	277	22	AA039502
3	277	51.7	277	22	AA039502
4	277	51.7	277	22	AA039502
5	260	48.5	260	22	AA014412
6	189	35.3	276	22	AA064506
7	136	25.4	190	22	AA089353
8	112	20.9	120	22	AB070328
9	45	8.4	277	22	AA014413
10	25	4.7	261	22	AA038694
11	25	4.7	261	22	AA073494
12	25	4.7	337	21	AA056582
13	25	4.7	337	23	AB041393
14	19	3.5	73	22	AB003475
15	19	3.5	73	22	AB012769
16	15	2.8	125	23	AA016594
17	15	2.8	260	22	AB062307
18	15	2.8	335	22	AA040480

ALIGNMENTS

RESULT 1	
AA041288	
ID	AA041288 standard; Protein; 296 AA.
AC	AA041288;
XX	
DT	22-OCT-2001 (first entry)
DE	Human polypeptide SEQ ID NO 6219.
XX	
KW	Human; noctropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW	peripheral nervous system; neuropathy; central nervous system; CNS;
KW	Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW	amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemocarcin;
KW	chemokine; thrombolytic; drug screening; arthritis; inflammation;
KW	leukemia.
XX	
OS	Homo sapiens.
XX	
PN	WO200153312-A1.
XX	
PD	26-UTL-2001.
XX	
PF	26-DEC-2000; 2000WO-US34263.
XX	
PR	21-JAN-2000; 2000US-0488725.
PR	05-APR-2000; 2000US-0552317.
PR	09-UTL-2000; 2000US-0598042.
PR	19-UTL-2000; 2000US-0620312.
PR	03-AUG-2000; 2000US-0653450.
PR	14-SEP-2000; 2000US-0662191.
PR	19-OCT-2000; 2000US-0693036.
PR	29-NOV-2000; 2000US-0727344.
XX	
PA	(HYSE-) HYSEQ INC.
XX	
PI	Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI	Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI	Zhao QA, Zhou F, Goodrich R, Dermanac RT;
XX	
DR	WPI; 2001-442253/47.
DR	N-PSDB; AA160444.
XX	
PT	Novel nucleic acids and polypeptides, useful for treating disorders
	such as central nervous system injuries -

XX Example 2; SEQ ID NO 6219; 10078pp; English.

PS The invention relates to human nucleic acids (AA157798-AA161369) and

XX the encoded polypeptides (AAM38642-AAM42213) with neurotropic,

CC immunosuppressant and cytoskeletal activity. The polynucleotides are useful

CC in gene therapy. A composition containing a polypeptide or polynucleotide

CC of the invention may be used to treat diseases of the peripheral nervous

CC system, such as peripheral nervous injuries, peripheral neuropathy and

CC localised neuropathies and central nervous system diseases, such as

CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic

CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the

CC utilisation of the activities such as: Immune system suppression,

CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic

CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,

CC assays for receptor activity, arthritis and inflammation, Leukemias and

CC C.N.S disorders.

CC Note: The sequence data for this patent did not form part of the printed

CC specification.

XX

XX Sequence 296 AA:

Alignment Scores:

Pred. No.:	Length:	Score:	Matches:	Conservative:	Best Local Similarity:	Mismatches:	Indels:	Query Match:	Gaps:
1,3e-295	296	296.00	296	0	100.00%	0	0	55.22%	0

US-09-896-522-1 (1-1624) x AAM41288 (1-296)

QY 37 GCCCGGGGGGGGAAAGGGGGGGGGGACCCCGATGCCGGAGGCGGAGCCCGAGATG 96

Db 1 AAlaGlyProGlyGluGlyAlaGlyAlaGlyThrArgCysAlaGlyAlaGluAlaGluMet 20

QY 97 GCTTCGGCGGGAGGCGAAGACTGCGAGAGCCCGCGCGCGAGCGGACCGTCCGACACAG 156

Db 21 AlaSerAlaGlyGlyGluAspCysGluSerProAlaProGluAlaAspArgProHisGln 40

QY 157 CGGCGCTTCTGTATGAGGGGTGAGCGCGGCACTCCAGCGGGAAGTGCAGCGTGTGAG 216

Db 41 ArgProHeuLeuIleGlyValSerGlyGlyThrAlaSerGlyLysSerThrValCysGlu 60

QY 217 AAGATCATGAGTGTCTGGGACAGAAACGAGTGGAAACGCGGACCGGAGAGTGTATC 276

Db 61 LysIleMetGluLeuLeuGlyGlnAsnGluValGluGlnArgGlnArgLysValIle 80

QY 277 CTGAGCCAGGACGAGTCTCAAGAGTCTTGAAGGCGAGAGCGAAGCGACCTTGAAA 336

Db 81 LeuSerGlnAspArgPheIleValLeuThrAlaGluGlnLysValAlaLysAlaLeuLys 100

QY 337 GGACAGTAACTTTTGAACATCCAGATGCCCTTTGATATGATTTGATGACAGGACTCTG 396

Db 101 GlyIleThrAsnPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThrLeu 120

QY 397 AAGAACAATCCGGGAGCGGAAACGCTGAGGTCGCCACTATGATTTTTCGACACACTCA 456

Db 121 LysAsnIleValGluGlyLysThrValGluValProThrArgPheValThrHisSer 140

QY 457 AGGTACAGAGACACGAGTGTCTACCGTCCGAGCGAGTGTCTGTGAGGAGCATCTTG 516

Db 141 ArgLeuProGluThrValValIleValIleProAlaAspValValLeuPheGluGlyIleLeu 160

QY 517 GTGTTCTACAGCCGAGAGATCCGAGACATGTTCCAGCTGCGCTCTTCTGTGACACCGAC 576

Db 161 ValPheTyrSerGlnGluIleArgAspMetPheHisLeuArgLeuPheValAspThrAsp 180

QY 577 TCCGACGTCAAGCTCTCCGAAGAGTTCCTCGGAGCGTCCGCCGAGGAGGACCTTGAG 636

Db 181 SerAspValArgLeuSerArgArgValLeuArgAspValArgArgGlyArgAspLeuGlu 200

QY 637 CAGATTCTGACGACGACACACACTTGGTGAAGCGCGGCTTCGAGGAGTTCGCTCGCG 696

Db 201 GlnIleLeuThrGlnIleThrThrPheValIleProAlaPheGluLubHeCysLeuPro 220

QY 697 ACAAGAAGTAGCCGATGATCATCCACGAGAGTGCACAAATGTTGGTCATCAAC 756

Db 221 ThrIleTyrThrAlaAspValIleIleProArgGlyValAspAsnMetValAlaIleAsn 240

QY 757 CTGATGTCGAGACACATCCAGGACATTTGAAAGTGGACATCGCAATGGGACCGAGGA 816

Db 241 LeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysLysTrpHisArgGly 260

QY 817 GGATCCAAATGGCGGAGCTACAAAGCGACCTTTTTCGAGCGGAGGAGCCAGCTGGAGATG 876

Db 261 GlySerAsnIleValArgSerTyrLysArgThrPheSerGluProGlyAspHisProGlyMet 280

QY 877 CTGACCTCTGGCAACGCTCACATTTGGAGTCCAGACGACACCCAC 924

Db 281 LeuThrSerGlyLysArgSerHisLeuGluSerSerArgProHis 296

RESULT 2

AAM39502

ID AAM39502 standard; Protein; 277 AA.

XX

XX AAM39502;

AC

XX

DT 22-OCT-2001 (first entry)

XX

DE Human polypeptide SEQ ID NO 2647.

XX

XX Human; neurotropic; immunosuppressant; cytoskeletal; gene therapy; cancer;

KM peripheral nervous system; neuropathy; central nervous system; CNS;

KM Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;

KM amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;

KM chemokinetic; thrombolytic; drug screening; arthritis; inflammation;

KM Leukemia.

XX

XX Homo sapiens.

XX

XX NO200153312-A1.

XX

PD 26-JUL-2001.

XX

XX 26-DEC-2000; 2000MO-US34263.

XX

PR 21-JAN-2000; 2000US-0488725.

PR 25-APR-2000; 2000US-0523217.

PR 09-JUL-2000; 2000US-0598042.

PR 19-JUL-2000; 2000US-0620312.

PR 03-AUG-2000; 2000US-0653450.

PR 14-SEP-2000; 2000US-0662191.

PR 19-OCT-2000; 2000US-0693036.

PR 29-NOV-2000; 2000US-0727344.

XX

XX (HYSE-) HYSEQ INC.

PA

XX

PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;

PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;

PI Zhao QA, Zhou F, Goodrich R, Drmanac RT;

XX

XX WPI; 2001-442253/47.

DR N-PDB; AA18658.

XX

XX Novel nucleic acids and polypeptides, useful for treating disorders

PT such as central nervous system injuries -

XX

XX Example 4; SEQ ID NO 2647; 10078pp; English.

XX

CC The invention relates to human nucleic acids (AA157798-AA161369) and

CC the encoded polypeptides (AAM38642-AAM42213) with neurotropic,

CC immunosuppressant and cytoskeletal activity. The polynucleotides are useful

CC in gene therapy. A composition containing a polypeptide or polynucleotide

CC of the invention may be used to treat diseases of the peripheral nervous

CC system, such as peripheral nervous injuries, peripheral neuropathy and

Score:	277.00	Matches:	277
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	51.68%	Indels:	0
DB:	22	Gaps:	0
US-09-896-522-1 (1-1624) x AAB93941 (1-277)			

[illegible][illegible]

DR WPI; 2002-140091/18.
 DR N-PSDB; AAD27186.
 XX
 PT New isolated human uridine kinase family polypeptide 57658, useful for
 PT treating hematopoietic neoplastic disorders and disorders of neurons,
 PT heart and blood vessels
 XX
 PS Claim 4; Fig 1a; 103pp; English.
 XX
 CC The patent discloses human uridine kinase-like polypeptides, designated
 CC 57658 and polynucleotides encoding such proteins. 57658 DNAs and proteins
 CC are useful for developing diagnostic and therapeutic agents for 57658-
 CC mediated or related disorders such as haematopoietic neoplastic disorders
 CC (e.g. leukemias), haematoid disorders (e.g. psoriasis, dermatitis,
 CC diabetes mellitus, rheumatoid arthritis, disorders of bone (e.g. rickets,
 CC osteoporosis), heart (e.g. sarcoma, myocardial infarction), blood vessels
 CC (e.g. hypertension, atherosclerosis) and neurons (e.g. Huntington's
 CC disease, Alzheimer's disease, Parkinson's disease). 57658 sequences or
 CC their antibodies are useful in screening assays, detection assays (e.g.
 CC forensic biology) and predictive medicine (e.g. diagnostic assays,
 CC prognostic assays, and monitoring clinical trials and pharmacogenomics).
 CC They are useful as reagents for diagnosing and treating 57658-mediated
 CC disorders. 57658 DNAs are useful in gene therapy, in chromosome mapping,
 CC to identify an individual from a minute biological sample (tissue typing)
 CC and to aid in forensic identification of the biological sample. The
 CC present sequence is human 57658 protein.

XX Sequence 277 AA;

Alignment Scores:

Pred. No.:	4,91e-276	Length:	277
Score:	277.00	Matches:	277
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	51.68%	Indels:	0
DB:	23	Gaps:	0

US-09-896-522-1 (1-1624) x AAB16592 (1-277)

QY 94 ATGGCTTCGGGAGGCGAAGACTGCGAGAGCCCGCGGAGCCGACCTCCGCAC 153
 DB 1 MetAlaSerAlaGlyGluAspCysGluSerProAlaProGluAlaAspArgProHis 20
 QY 154 CACGGCCCTTCCTGATAGGGGTAGCGGCGACCTCCGCGGAAGTGCACCTGTGT 213
 DB 21 GlnArgProHeuLeuIleGlyValSerGlyGlyThrAlaSerGlyLysSerThrValCys 40
 QY 214 GAGAAATCATGAGTTGCTGAGCAGAAAGAGGTGGAGCGGCGAGCGAGAGTGTGTC 273
 DB 41 GluLysIleMetCtluLeuLeuGlyGlnAsnGluValGluGlnArgGlnArgLysVal 60
 QY 274 ATCTTAGAGCAGAGACAGTTTCTACAAAGTCTTGAACGGCAGAGCAAGAGCCCTTG 333
 DB 61 IleLeuSerGlnAspArgPheTyrLysValLeuThrAlaGluGlnLysAlaLysAlaLeu 80
 QY 334 AAGGACAGTAACTTTTGCACATCCAGATGCTTGTATATGATTGTGACAGACACT 333
 DB 81 LysGlyGlnIleTyrAsnPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThr 100
 QY 394 CTGAAGAACATCGTGGAGGCGCAAAACGATGAGGTGGCGGACCTATGATTGTCACAC 453
 DB 101 LeuLysAsnIleValGluGlyLysThrValGluValProThrLysAspPheValThrHis 120
 QY 454 TCAGAGTTACAGAGACACGCGTGTCTACCTCGCGAGCGTGTCTGTTTGAAGGCATC 513
 DB 121 SerArgLeuProGluThrThrValValTyrProAlaAspValValLeuPheGluGlyIle 140
 QY 514 TTGGTGTTCACAGCCAGAGATCCGGAGATGTCCACCTGGCGCTTCGTGGACACC 573
 DB 141 LeuValPheTyrSerGlnGlnIleArgAspMetPheHisLeuLysArgLeuPheValAspThr 160
 QY 574 GACTCCGACGTCAGGCTGTCTCGAAGAGTCTCCGGAGCCTGGCCGAGGAGGAGCACTG 633

DB 161 AspSerAspValArgLeuSerArgArgValLeuArgAspValArgArgGlyArgAspLeu 180
 QY 634 GAGCAGATTCTGACCGAGTACACCACTTCGTGGAAGCCGGCTTCGAGAGTCTCGCTG 693
 DB 181 GluGlnIleLeuThrIleTyrThrThrPheValLysProAlaPheGluGluPheCysLeu 200
 QY 694 CCGCAAAAGAGATGCCGATGATCATCCACGAGAGTGGACATATGTTGGCATC 753
 DB 201 ProThrLysAlaTyrLysAlaAspValIleIleProArgGlyValAlaAspAsnMetValAlaIle 220
 QY 754 AACCTGATCGTGACGACATCCAGACATTTGAAATGTGACATTCGAAATGGACCGA 813
 DB 221 AsnIleIleValGlnIleHisIleGlnAspIleLeuAsnGlyAspIleCysLysTrpHisArg 240
 QY 814 GAGGGTCCAAATGGGCGGAGCTTCAAGCGGACCTTTTTCGAGCCAGGAGGACCCCTGG 873
 DB 241 GlyGlySerAsnLysArgSerTyrLysAspGlnPheSerGluProLysAspHisProGly 260
 QY 874 ATGCTGACCTTGGCAAGGATCACATTTGAGTCCAGACGAGACCCAC 924
 DB 261 MetLeuThrSerGlyLysArgSerHisLeuGluSerSerArgProHis 277
 RESULT 5
 ID AA014412
 AA014412 standard; Protein; 260 AA.
 XX
 AC AA014412;
 XX
 DT 02-MAY-2002 (first entry)
 XX
 DE Protein of a human uridine kinase (UDK).
 XX
 KW Human; uridine kinase; diagnostic assay; mutation detection; UDK;
 KW probe; chromosome localization study; tissue expression; gene therapy;
 KW antibody; vaccine; human ovarian cancer; immunological disorder;
 KW human colon carcinoma; immunogen.
 XX
 OS Homo sapiens.
 XX
 PN WO200172963-A2.
 XX
 PD 04-OCT-2001.
 XX
 PF 27-MAR-2001; 2001WO-US09663.
 XX
 PR 27-MAR-2000; 2000US-0536647.
 XX
 PA (SMIK) SMITHKLINE BEECHAM CORP.
 PI Ho YS, Johnson RK;
 XX
 DR WPI; 2001-626259/72.
 DR N-PSDB; AAK98735.
 XX
 PT Novel human uridine kinase polypeptides useful for treating cancers,
 PT and to identify agonists and antagonists of the polypeptide useful for
 PT treating conditions associated with uridine kinase imbalance
 PS Claim 3; Page 29-30; 31pp; English.
 XX
 CC The invention relates to newly identified human uridine kinase (UDK)
 CC polypeptides and polynucleotides and methods for producing such
 CC polypeptides by recombinant techniques. Also disclosed in the invention
 CC are methods for utilizing uridine kinase polypeptides and polynucleotides
 CC in diagnostic assays. The polynucleotides and polypeptides of the
 CC invention may be used as diagnostic reagents by detecting mutations in an
 CC associated gene. An array of oligonucleotide probes comprising the
 CC uridine kinase polynucleotide sequence or fragments thereof can be
 CC constructed to conduct efficient screening of genetic mutations, for
 CC example. Detection of abnormally decreased or increased levels of
 CC polypeptide or mRNA expression may also be used for diagnosing or
 CC determining susceptibility of a subject to a disease of the invention.
 CC The polynucleotide sequences of the invention can be used for chromosome

CC localisation studies and tissue expression studies. The polypeptides of
CC the invention or fragments thereof may be used as immunogens to produce
CC antibodies. These antibodies may be employed to isolate or identify
CC clones expressing the polypeptide. The polypeptides and polynucleotides
CC of the invention can be used as a vaccine or in gene therapy to treat
CC diseases such as human ovarian cancer, human colon carcinomas, and
CC immunological disorders. This sequence represents the protein of a human
CC uridine kinase of the invention.

XX Sequence 260 AA;

Alignment Scores:
Pred. No.: 1.62e-258 Length: 260
Score: 260.00 Matches: 260
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 48.51% Indels: 0
DB: 22 Gaps: 0

US-09-896-522-1 (1-1624) x AAO14412 (1-260)

QY 94 ATGGCTTCGGCGGAGGCGAAGACTGGGAGAGCCCGCGCGGAGCCGACCTCCGAC 153
DB 1 MetalaSerAlaGlyGlyLysAspGlySerProAlaProGluAlaAspArgProHis 20
QY 154 CAGCGGCTTCCTGATAGGGGTGAGCGGCGGCACTGCGGAGGAAAGTCAGCCGTGT 213
DB 21 GlnArgProPheLeuIleGlyValSerGlyIleThrAlaSerGlyLysSerThrValCys 40
QY 214 GAGAAATCATGATGCTGCTGGACAGAACAGAGGTGAAACGCGGCGAGCGAGCTGTC 273
DB 41 GlnLysIleMetGlnLeuIleGlnIleValGlnAsnGlnValGlnGlnArgIleValVal 60
QY 274 ATCTGAGCCGAGGACAGGTTCTACAGGTCCTGAOGGAGGAGGAGGCGCAAGGCTTG 333
DB 61 IleuSerGlnAspArgPheTyLysValLeuThrAlaGlnGlnLysAlaLysAlaLeu 80
QY 334 AAAGACAGTAACTTTTGACATCCAGATCCCTTTGATATGATTGATGACAGAGCT 393
DB 81 LysGlyGlnTyLysPheAspPheHisProAlaPheAspAsnAspLeuMetHisArgThr 100
QY 394 CTGAAGAAATCGTGGAGGGGAAACGGTGGAGGTCCGACCTTGTGATTTGTGACAC 453
DB 101 LeuLysAsnIleValGlnGlyLysThrValGlnValProThrTyLysAspPheValThrHis 120
QY 454 TCAAGTTACAGAGACACAGGAGTGTACCTCCGCGGAGCGGTTCTGTTGAGGGGATC 513
DB 121 SerArgLeuProGlnThrValValValTyLysProAlaAspValValLeuPheGlnGlyLe 140
QY 514 TTGGTGTCTTACAGCAGAGATCCGGGACATGTTCCACCTGCGCTTCTGTGACACC 573
DB 141 LeuValPheTyLysSerGlnGlnIleArgAspPheHisLeuArgLeuPheValAspThr 160
QY 574 GACTCCGACGTCAGGCTGTCTCGAAGAGTTTCCGGGACGTCGCGGAGGAGGACCTG 633
DB 161 AspSerAspValArgLeuSerArgValLeuArgAspValArgArgGlyArgAspLeu 180
QY 634 GAGAGATTCTGAGCAGTACACACCTGCTGAAGCGGCTTCGAGGAGTTCTGCTG 693
DB 181 GlnGlnIleLeuThrGlnTyLysThrPheValLysProAlaPheGlnGlnPheCysLeu 200
QY 694 CCGAAGAAAGTATGCGCATGTGATCATCCAGAGAGTGAACAATATGTTGCTCATC 753
DB 201 ProThrLysLysTyLysAlaAspValIleIleProArgGlyValAlaAspMetValAlaIle 220
QY 754 AACTGATCTGCGACGACATCCAGGACATTTGATGTTGATCATCTGCAATATGGACCGA 813
DB 221 AsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysLysThrPheArg 240
QY 814 GAGAGGTCCATGGGCGGAGCTAACAGCGGACCTTTCTGAGCGGAGGAGCAACCTGGG 873
DB 241 GlyGlySerAsnGlyArgSerTyLysArgThrPheSerGlnProGlyAspHisProGly 260

RESULT 6
AAG64506
ID AAG64506 standard; Protein; 276 AA.

XX AAG64506;

AC AAG64506;

DT 02-OCT-2001 (first entry)

XX Human uridine kinase.

XX Human; uridine kinase, UK.

XX Homo sapiens.

XX CN1287172-A.

XX 14-MAR-2001.

XX 07-SEP-1999; 99CN-0118818.

XX 07-SEP-1999; 99CN-0118818.

XX (YTFU-) UNITV FUDAN.

XX Yu L, Zhao Y, Zhang H;

XX WPI: 2001-409529/44.

XX N-PSDB; AAH75355.

XX Human uridine kinase and its coding sequence, preparation and

XX application -

XX Claim 2; Page 15-16 (Disclosure); 20pp; Chinese.

XX The invention relates to human uridine kinase (UK).

XX Sequence 276 AA;

Alignment Scores:
Pred. No.: 2.28e-185 Length: 276
Score: 189.00 Matches: 189
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 35.26% Indels: 0
DB: 22 Gaps: 0

US-09-896-522-1 (1-1624) x AAG64506 (1-276)

QY 265 AAGGTGTCATCTGAGCCAGGACAGGTTCTACAGGTCCTGACGACAGAAAGCC 324
DB 57 LysValValIleLeuSerGlnAspArgPheTyLysValLeuThrAlaGlnGlnLysAla 76
QY 325 AAGGCTTGAAGAGACATCAATTTTGACATCCAGATGCTTGTATGATTTGATG 384
DB 77 LysAlaLeuLysGlyGlnTyLysPheAspPheHisProAlaPheAspAsnAspLeuMet 96
QY 385 CACAGAGCTCTGAAGAAATCGTGGAGGCAAAACGTTGAGAGTCCGACCTATGATTTT 444
DB 97 HisArgThrLeuLysAsnIleValGlnGlyLysThrValGlnValProThrTyLysPhe 116
QY 445 GTGACACACTCAAGTTTACAGGACACAGAGTGTGTCACCTGCGGACGTGTTCTGTT 504
DB 117 ValThrHisSerArgLeuProGlnThrValValTyLysProAlaAspValValLeuPhe 136
QY 505 GAGGGCATCTGGTGTCTTACAGCCAGGAGATCCGGGACATGTTCCACCTGCGCTTTC 564
DB 137 GlnGlyIleLeuValPheTyLysSerGlnGlnIleArgAspPheHisLeuArgLeuPhe 156
QY 565 GTGACACGAGCTCCGAGCTGTCGAAAGGTTCTCCGGGACGTGCGGCGAGG 624
DB 157 ValAspThrAspSerAspValArgLeuSerArgValLeuArgAspValArgArgGly 176
QY 625 AGGAGCTTGAGAGATTTCTGACGACATCACACCTTCTGTGAAGCGGCTTTCGAGAG 684

Db 177 ArgAspLeuInGlnLeuThrInGlnTyr-ThrThrPheValLysProAlaPheGlu 196
 Qy 685 TTCTGCTGCGCCAAAGATATGCCGATGTATCTCCACGAGAGTGGACATATG 744
 Db 197 PheCysLeuProThrLysLysTyrAlaAspValIleIleProArgGlyValAspAsnMet 216
 Qy 745 GTTGCCATCAACCGATCGTGCAGACATCGAGACATTCGTGAATGGTGCATCTGCAA 804
 Db 217 ValAlaIleAsnLeuIleValGlnHisIleGlnHisPheLeuAsnGlyAspIleCysLys 236
 Qy 805 TGGCACCAGAGAGGTGCCAATGGCGG 831
 Db 237 TrpHisArgGlyGlySerAsnGlyArg 245
 RESULT 7
 ABB89353
 ID ABB89353 standard; Protein; 190 AA.
 AC ABB89353;
 XX 24-MAY-2002 (first entry)
 DE Human polypeptide SEQ ID NO 1729.
 XX
 KM Cytosolic; immunosuppressive; nootropic; neuroprotective; antiviral;
 KM antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor;
 KM valine; anticonvulsant; antibacterial; antifungal; antiparasitic;
 KM cardiac; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KM neurological disease; infection; human; secreted protein.
 XX
 OS Homo sapiens.
 PN WO200190304-A2.
 PD 29-NOV-2001.
 XX
 PF 18-MAY-2001; 2001MO-US16450.
 XX
 PR 19-MAY-2000; 2000US-205515P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Birse CE, Rosen CA;
 XX WPI; 2002-122018/16.
 DR N-PSDB; ABL89762.
 XX
 PT Novel 1405 isolated polypeptides, useful for diagnosis, treatment and
 PT prevention of neural, immune system, muscular, reproductive,
 PT gastrointestinal, pulmonary, cardiovascular, renal and proliferative
 PT disorders -
 XX
 PS Claim 11; SEQ ID NO 1729; 2081tp + Sequence listing; English.
 XX
 CC The invention relates to novel genes (ABL89449-ABL90853) and proteins
 CC (ABB89040-ABB90444) useful for preventing, treating or ameliorating
 CC medical conditions e.g. by protein or gene therapy. The genes are
 CC isolated from a range of human tissues disclosed in the specification.
 CC The nucleic acids, proteins, antibodies and (ant)agonists are useful
 CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
 CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
 CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
 CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
 CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
 CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;
 CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
 CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
 CC and parasitic infections.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 190 AA;
 Alignment Scores:
 Pred. No.: 9,58e-131 Length: 190
 Score: 136.00 Matches: 136
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 25,37% Indels: 0
 DB: 23 Gaps: 0
 US-09-896-522-1 (1-1624) x ABB89353 (1-190)
 Qy 193 AGCGGGAGTCGACCGCTGTGTGAGAGATCATGGAATTGCTGGACAGAACGAGTGGAA 252
 Db 34 SerGlyLysSerThrValGlyGlnLysIleMetGlnLeuLeuGlnAsnGlnValGln 53
 Qy 253 CAGCGGCAGCGGAGAGGTGTGATCTCTGAGCCAGACAGAGTTCTACAGGTCCTGACGGCA 312
 Db 54 GlnArgGlnArgGlyValValIleLeuSerGlnAspArgPheTyrIleValLeuThrAla 73
 Qy 313 GAGCAGAGGCGCAAGGCTTGAAGAGACATACATTTTGAACCATCCAGATGCCCTTGAAT 372
 Db 74 GluGlnLysAlaIleValLeuLysGlnTyrAsnPheAspHisProAspAlaPheAsp 93
 Qy 373 AATGATTTGATGACACGACTCTGAGAAATCTGAGAGGCGCAAAACGGTGGAGTGGCGG 432
 Db 94 AsnAspLeuMetHisArgThrLeuLysAsnIleValGlnGlyLysThrValGlnValPro 113
 Qy 433 ACCTAGATTTTGTGACACACTCAAGTTACAGACAGACAGCGGTGCTTACCTGGCGGAC 492
 Db 114 ThrTyrAspPheAlaThrIleSerArgLeuProGlnThrThrValValTyrProAlaAsp 133
 Qy 493 GTGTTCTGTGTTGAGGGCATCTTGTGTTCTACAGCCAGAGATCCGGGACATGTTCCAC 552
 Db 134 ValValLeuPheGlnGlyIleLeuValPheTyrSerGlnGlnIleArgAspMetPheHis 153
 Qy 553 CTGCGGCTCTCTGCTGAGACACCGACTCCGACGTCGACGCTGTCGAGAA 600
 Db 154 LeuArgLeuPheValAspThrAspSerAspValArgLeuSerArgArg 169
 RESULT 8
 ABB70328
 ID ABB70328 standard; Protein; 120 AA.
 AC ABB70328;
 XX
 XX 21-OCT-2002 (first entry)
 DE Human MBDT protein Incyte ID No: LI:235557.12.orf2:2001JAN12.
 XX
 KM Human; molecule for disease detection and treatment; MBDT; cancer;
 KM cell proliferative disorder; arteriosclerosis; cirrhosis; asthma;
 KM autoimmune disorder; inflammatory disorder; Crohn's disease;
 KM multiple sclerosis; cytotoxic; antiarteriosclerotic; antiinflammatory;
 KM hepatotropic; immunosuppressive; antiasthmatic.
 XX
 OS Homo sapiens.
 PN WO200255738-A2.
 PD 18-JUL-2002.
 XX
 PD 09-JAN-2002; 2002MO-US01008.
 PF 12-JAN-2001; 2001US-261622P.
 XX 16-JAN-2001; 2001US-261865P.
 PR 17-JAN-2001; 2001US-262208P.
 PR 17-JAN-2001; 2001US-262209P.
 PR 17-JAN-2001; 2001US-262326P.
 PR 19-JAN-2001; 2001US-263063P.
 PR 19-JAN-2001; 2001US-263065P.
 PR 19-JAN-2001; 2001US-263329P.

XX	(INCY-)	INCYTE GENOMICS INC.
XX		
PI	Panzer SR,	Lincoln SE, Altus CM, Dufour GE, Hillman JL, Jones AJ,
PI	Dam TC,	Liu TP, Harris B, Flores V, Dafio A, Marwaha R, Chen AV;
PI	Chang SC,	Gerstlin EH, Peralta CH, David MH, Lewis SA;
XX		
DR	WPI; 2002-590679/63.	
DR	N-Psdb; ABS51801.	
PT	New disease detection and treatment molecule (MDR) polynucleotides and	
PT	polypeptides, useful in diagnosing, studying, preventing or treating	
PT	diseases associated with MDR expression, e.g. autoimmune or	
PT	inflammatory disorders -	
XX		
PS	Claim 27; Page 120; 129pp; English.	
XX		
CC	The present invention relates to the isolation of novel human	
CC	molecules for disease detection and treatment (MDR), and the	
CC	polynucleotide sequences (mdr) encoding them. The MDR polypeptides	
CC	may be used to screen for molecules that bind to, or are bound by the	
CC	encoded polypeptides, and to develop a transcript image of a tissue or	
CC	cell type. Probes comprising at least 20 nucleotides of the mdr	
CC	polynucleotide may be used to assess the toxicity of a test compound.	
CC	The MDR polypeptides and mdr polynucleotides are useful in the	
CC	diagnosis, study, prevention and treatment of diseases associated with	
CC	the expression of molecules for disease detection and treatment. Such	
CC	disorders include cell proliferative disorders (e.g. arteriosclerosis,	
CC	cirrhosis, or cancers), and autoimmune/inflammatory disorders	
CC	(e.g. asthma, Crohn's disease, or multiple sclerosis). The mdr	
CC	polynucleotides may also be used as molecule markers, in microarrays,	
CC	and in somatic or germline gene therapy. ABG70306-ABG70341 represent	
CC	the MDR proteins of the invention.	
XX		
XX		
SQ	Sequence	120 AA;
Alignment Scores:		
Pred. No.:	5.37e-106	Length: 120
Score:	112.00	Matches: 112
Percent Similarity:	100.00%	Conservative: 0
Best Local Similarity:	100.00%	Mismatches: 0
Query Match:	20.90%	Indels: 0
DB:	23	Gaps: 0
US-09-896-522-1 (1-1624) x ABG70328 (1-120)		
OY	361 GATGCCCTTTGATATGATTGGATGCACAGACTGTGAAGAATCGTGAGGCAAAACG	420
Db	5 Asplaphheaspndspneuspneuwethisatgtnhtleuylsaenillevalglugllylvsthr	24
OY	421 GTGAGAGTGGCCGACCTATGATTTTGTGACACACTCAAGTTTACGAGAACACGGTGGTC	480
Db	25 Valglivalprothrlyrzaphepnevalttrh1sserfargleuprogutThrValval1	44
OY	481 TACCCTCGGACGCTGGTCTGTGTTTGAAGGCACTTGGTGTTTACAGCAGACAATCCGG	540
Db	45 TyrProhlaaspyalvalleuphegiungllyllelevalphetrySergInulinlearg	64
OY	541 GACATGTTCCACCTCGCGCCTTTGTGACACACGCACTCCGACGTCGAGCTGTCTGAGA	600
Db	65 AspmetPhehisleuarlgueuphevalAseptrnaspserraspvalargLeuserArgarg	84
OY	601 GTTCTCCGGAGCGTGGCGGAGGAGGACCTGGAGCAGATTCTGACCGCACTACACACC	660
Db	85 ValleunargspyalrargargglyargabpbeuigiunllileuthrgIntyrtThr	1040
OY	661 TTGCTGAAGCCGGCCTTCGAGAGAGTTCTGCGCGG	696
Db	105 PheVallysrProhlaaphneglunglunPhesyLeupro	116
RESULT_9		
AAOI4413		
ID	AAOI4413 standard; Protein; 277 AA.	

```

XX AA014413;
AC
DT 02-MAY-2002 (first entry)
DE Protein relating to a human uridine kinase (UDK) of the invention.
XX
XX Human; uridine kinase; diagnostic assay; mutation detection; UDK;
KW probe; chromosome localisation study; tissue expression; gene therapy;
KW antibody; vaccine; human ovarian cancer; immunological disorder;
KW human colon carcinoma; immunogen.
XX
OS Unidentified.
XX
XX WO200172363-A2.
XX
XX 04-OCT-2001.
XX
XX 27-MAR-2001; 2001MO-US09663.
XX
XX 27-MAR-2000; 2000US-0536647.
XX
XX (SMIK ) SMITHKLINE BEECHAM CORP.
XX
XX Ho YS, Johnson RK;
XX
XX WPI; 2001-626259/72.
XX
XX
XX Novel human uridine kinase polypeptides useful for treating cancers,
PT and to identify agonists and antagonists of the polypeptide useful for
PT treating conditions associated with uridine kinase imbalance -
XX
XX Disclosure; Page 23; 31pp; English.
XX
XX The invention relates to newly identified human uridine kinase (UDK)
CC polypeptides and polynucleotides and methods for producing such
CC polypeptides by recombinant techniques. Also disclosed in the invention
CC are methods for utilising uridine kinase polypeptides and polynucleotides
CC in diagnostic assays. The polynucleotides and polypeptides of the
CC invention may be used as diagnostic reagents by detecting mutations in an
CC associated gene. An array of oligonucleotide probes comprising the
CC uridine kinase polynucleotide sequence or fragments thereof can be
CC constructed to conduct efficient screening of genetic mutations, for
CC example. Detection of abnormally decreased or increased levels of
CC polypeptide or mRNA expression may also be used for diagnosing or
CC determining susceptibility of a subject to a disease of the invention.
CC The polynucleotide sequences of the invention can be used for chromosome
CC localisation studies and tissue expression studies. The polypeptides of
CC the invention or fragments thereof may be used as immunogens to produce
CC antibodies. These antibodies may be employed to isolate or identify
CC clones expressing the polypeptide. The polypeptides and polynucleotides
CC of the invention can be used as a vaccine or in gene therapy to treat
CC diseases such as human ovarian cancer, human colon carcinomas, and
CC immunological disorders. This sequence represents the protein relating to
CC a human uridine kinase (UDK) of the invention.
CC
CC NOTE: The present sequence is stated as being the same as that shown as
CC SEQ ID NO: 2 in the sequence listing of the specification. However, the
CC sequences differ.
XX
XX
XX Sequence 277 AA;
XX
XX Alignment Scores:
XX Pred. No.: 5.26e-37 Length: 277
XX Score: 45.00 Matches: 45
XX Percent Similarity: 100.00% Conservative: 0
XX Best Local Similarity: 100.00% Mismatches: 0
XX Query Match: 8.40% Indels: 0
XX Gaps: 0
XX
XX US-09-896-522-1 (1-1624) x AA014413 (1-277)
XX
XX 391 ACTCTGAAGAACATGCTGGAGGGCAAACCGTGAGTGCCGACCTTATGATTTTGACAA 450

```


Db 100 ThrlautysanileValGlulylLeThrValGlulValProThrTyraaspheValThr 119
QY 451 CACTCAGGTTTACCAGACGACCGTGTCTACCTCGGACGCGTTCGTTGTTGAGGC 510
Db 120 HisserTrgleupProGluTrThrValValTyProAlaSpValValLeupheGluGly 139
QY 511 ATCTGGTGTCTTAC 525
Db 140 lleleuValphetyr 144
RESULT 10
AAM38694
ID AAM38694 standard; Protein; 261 AA.
XX
AC AAM38694;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polypeptide SEQ ID NO 1839.
XX
KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.
XX
OS Homo sapiens.
XX
PN WO200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US34263.
XX
XX 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
XX
XX (HYSE-) HYSEQ INC.
PA
XX
PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
DR WPI; 2001-442253/47.
DR N-PSDB; AAI57850.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders
XX such as central nervous system injuries -
PS
PS Example 3; SEQ ID NO 1839; 10078bp; English.
XX
XX The invention relates to human nucleic acids (AA157798-AA161369) and
XX the encoded polypeptides (AAM38642-AAM42213) with nootropic,
XX immunosuppressant and cytoskeletal activity. The polynucleotides are useful
XX in gene therapy. A composition containing a polypeptide or polynucleotide
XX of the invention may be used to treat diseases of the peripheral nervous
XX system, such as peripheral nervous injuries, peripheral neuropathy and
XX localised neuropathies and central nervous system diseases, such as
XX Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
XX lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
XX utilisation of the activities such as: Immune system suppression,
XX Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
XX and thrombolytic activity, cancer diagnosis and therapy, drug screening,
XX assays for receptor activity, arthritis and inflammation, leukaemias and
XX C.N.S disorders.
CC
CC Note: The sequence data for this patent did not form part of the printed

CC specification.
XX
SQ Sequence 261 AA;
XX
Alignment Scores:
Pred. No.: 2,14e-16 Length: 261
Score: 25.00 Matches: 25
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 4.66% Indels: 0
DB: 22 Gaps: 0
US-09-896-522-1 (1-1624) x AAM38694 (1-261)
QY 658 ACCTTGCTGAAGCGCGCTTCGAGAGTTCGCTCGCGACAAAGATGCGCATGTC 717
Db 187 ThrPheVallyseProAlaPheGluGluGluPheCysLeuProThrTyraAlaSpVal 206
QY 718 ATCATCCACAGAGGA 732
Db 207 llelleProArggly 211
RESULT 11
AAB73494
ID AAB73494 standard; Protein; 261 AA.
XX
AC AAB73494;
XX
XX 31-JUL-2001 (first entry)
XX
DE Human transferase HTFS-1, SEQ ID NO:1.
XX
KW Human transferase; HTFS; agonist; antagonist; cellular signalling;
KW proliferation; cell proliferative disorder; immune disorder;
KW atherosclerosis; hepatitis; psoriasis; cancer; tumour;
KW inflammation; AIDS; Addison's disease; allergy; asthma; anaemia;
KW cirrhosis; Crohn's disease; atopic dermatitis; diabetes mellitus;
KW multiple sclerosis; rheumatoid arthritis; pancreatitis;
KW systemic lupus erythematosus; thrombocytopenia; ulcerative colitis;
KW haemodialysis; extracorporeal circulation; trauma; transgenic animal;
KW gene therapy; drug screening.
XX
XX Homo sapiens.
OS
XX
PN WO200132888-A2.
XX
PD 10-MAY-2001.
XX
PF 02-NOV-2000; 2000WO-US30485.
XX
XX 04-NOV-1999; 99US-0163595.
XX
XX (INCY-) INCYTE GENOMICS INC.
PA
XX
PI Tang YT, Yue H, Hillman JL, Lal P, Bandman O, Patterson C;
PI Shih LL, Azimzai Y, Lu DAM, Baughn MR;
XX
DR WPI; 2001-328796/34.
DR N-PSDB; AAB73801.
XX
PT Human transferase polypeptides and polynucleotides useful for
XX diagnosis, prevention and treatment of cell proliferative and immune
XX system disorders and for identifying agonists and antagonists -
XX
XX Claim 1; Page 103-104; 157pp; English.
XX
XX Sequences AAB73494-AAB73535 represent novel human transferase proteins
XX HTFS-1 to HTFS-42, and sequences AAB73801-AA723842 represent cDNAs
XX encoding them. The proteins play important roles in the regulation of
XX cellular signalling and proliferation. The HTFS proteins are useful for
XX screening compounds for their effectiveness as agonists or antagonists of
XX transferase activity, or for compounds that specifically bind to an HTFS
XX protein or which modulates the activity of an HTFS protein.

CC Pharmaceutical compositions comprising an HTFS protein, HTFS
CC agonist or antagonist, or genetic construct encoding an HTFS
CC protein are useful for treating a disease or condition associated
CC with decreased or increased expression of functional HTFS. Disorders
CC which may be treated using such compositions include cell proliferative
CC disorders and immune disorders. For example, diseases which may be
CC treated include atherosclerosis, hepatitis, psoriasis, cancers (including
CC breast, bladder, bone marrow, brain and uterus cancer), inflammation,
CC AIDS, Addison's disease, allergies, asthma, anaemia, cirrhosis, Crohn's
CC disease, atopic dermatitis, diabetes mellitus, multiple sclerosis,
CC rheumatoid arthritis, pancreatitis, systemic lupus erythematosus,
CC thrombocytopenia, and ulcerative colitis. They may also be used to treat
CC complications of cancer, haemodialysis, extracorporeal circulation,
CC leukaemia and haematopoietic cancer, including lymphoma, leukaemia and
CC myeloma. Polynucleotides encoding HTFS proteins are useful for creating
CC transgenic animals to model human diseases, for diagnostic purposes and
CC to generate hybridisation probes useful in mapping the naturally
CC occurring genomic sequences. HTFS, and its catalytic or immunogenic
CC fragments are useful for screening libraries of compounds in a variety of
CC drug screening techniques. Antibodies which specifically bind HTFS may be
CC used for the diagnosis of disorders associated with the expression of
CC HTFS, or in assays to monitor patients being treated with HTFS or
CC agonists, antagonists or inhibitors of HTFS. The present sequence
CC represents an HTFS protein of the invention.

XX Sequence 261 AA;

Alignment Scores:

Pred. No.:	2.14e-16	Length:	261
Score:	25.00	Matches:	25
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	4.66%	Indels:	0
DB:	22	Gaps:	0

US-09-896-522-1 (1-1624) x AAB73494 (1-261)

QY 658 ACCTTGCGAAGCGGCGCTTCGAGAGATTCTGCTGCCGACAAAGATATGCCGATGTG 717
Db 187 ThrpheVallyspProAlaPheGluGluPheCysLeuProThrlyblySTyAlaAspVal 206

QY 718 ATCATCCACGAGAGA 732
Db 207 IleIleProArgGly 211

RESULT 12

AAB56582 ID AAB56582 standard; Protein, 337 AA.

XX AC AAB56582;

XX DT 13-MAR-2001 (first entry)

DE Human prostate cancer antigen protein sequence SEQ ID NO:1160.

XX Human, prostate cancer; prostate cancer antigen; detection; diagnosis;
XX neuroprotective; cytoskeletal; cardiovascular; immunomodulatory; muscular;
XX vulvovaginal; gastrointestinal; nephrotoxic; antineoplastic; gynaecological;
XX antibacterial; gene therapy; neural; immune; reproductive; renal;
XX gastrointestinal; pulmonary; cardiovascular; proliferative disorder;
XX wound; infectious disease.

XX OS Homo sapiens.

XX PN W0200055174-A1.

XX PD 21-SEP-2000.

XX PF 08-MAR-2000; 2000WO-US05988.

XX PR 12-MAR-1999; 99US-0124270.

XX PA (HUMA-) HUMAN GENOME SCI INC.

PA (ROSE/) ROSEN C A.
XX Rosen CA, Ruben SM;
XX WPI, 2000-567513/55.
DR N-PSDB; AAF15785.

XX Prostate cancer associated gene sequences, referred to as prostate
XX cancer antigens, useful for treatment, prevention, and diagnosis of
XX disorders such as prostate cancer -

XX Claim 11; Page 1566-1567; 2338pp; English.

XX AAF15566 to AAF16505 encode the human prostate cancer associated
XX proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.
XX The prostate cancer antigens can have neuroprotective, cytoskeletal,
XX cardioprotective, immunomodulatory, muscular, vulvovaginal, gastrointestinal,
XX nephrotoxic, antineoplastic, gynaecological and antibacterial activities,
XX and can be used in gene therapy. The prostate cancer antigen
XX polynucleotides may be used for detection of prostate cancer, chromosome
XX identification, as chromosome markers, and for numerous other diagnostic
XX or research purposes. The prostate cancer antigens may be used to treat
XX disorders such as neural, immune, muscular, reproductive,
XX gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
XX disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to
XX AAB57303 represent sequences used in the exemplification of the present
XX invention.

XX Sequence 337 AA;

Alignment Scores:			
Pred. No.:	2.08e-16	Length:	337
Score:	25.00	Matches:	25
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	4.66%	Indels:	0
DB:	21	Gaps:	0

US-09-896-522-1 (1-1624) x AAB56582 (1-337)

QY 658 ACCTTGCGAAGCGGCGCTTCGAGAGATTCTGCTGCCGACAAAGATATGCCGATGTG 717
Db 263 ThrpheVallyspProAlaPheGluGluPheCysLeuProThrlyblySTyAlaAspVal 282

QY 718 ATCATCCACGAGAGA 732
Db 283 IleIleProArgGly 287

RESULT 13

ABP41393 ID ABP41393 standard; Protein, 337 AA.

XX AC ABP41393;

XX DT 22-AUG-2002 (first entry)

DE Human ovarian antigen HOBLP29, SEQ ID NO:2525.

XX Human, ovarian antigen; ovary; ovarian; breast; cancer; tumour;
XX ovarian cancer; breast cancer; tumour; reproductive system disorder;
XX infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
XX PCOS; ovarian cyst; dysmenorrhoea; endocrine disorder; infection;
XX inflammatory condition; immune disorder; blood disorder;
XX cardiovascular disorder; respiratory disorder; neurological disorder;
XX gastrointestinal disorder; urinary system disorder; drug screening;
XX gene therapy; chromosome mapping; forensic analysis;
XX antibody preparation; cytoskeletal; immunomodulatory; neuroprotective;
XX antiinflammatory; gynaecological; reproductive; chromosome 1p32.

XX OS Homo sapiens.

XX PN W0200200677-A1.

PR 12-SEP-2000; 2000US-0231966.
PR 14-SEP-2000; 2000US-0233397.
PR 14-SEP-2000; 2000US-0233398.
PR 14-SEP-2000; 2000US-0233399.
PR 14-SEP-2000; 2000US-0233400.
PR 14-SEP-2000; 2000US-0233401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235835.
PR 27-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236803.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.

PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-451937/48.
XX N-PSDB; AAL35057.
XX
XX
XX Isolated polypeptide for treating, preventing and/ or prognosing
XX disorders related to the musculoskeletal system including
XX musculoskeletal cancers and also for testing and detection e.g.
XX diagnosis -
XX
XX Claim 11; SEQ ID NO 1422; 781bp + Sequence listing; English.
XX
XX The invention relates to novel genes (AAL34669-AAL37666) and proteins
XX (ABB03087-ABB04109) associated with the musculoskeletal system useful
XX for preventing, treating or ameliorating medical conditions e.g. by
XX protein or gene therapy. The genes are isolated from a range of human
XX tissues disclosed in the specification. The nucleic acids, proteins,
XX antibodies and (ant)agonists are useful in the diagnosis, treatment
XX and prevention of: (a) cancer, e.g. breast and ovarian cancer and
XX other cancers of the adrenal gland, bone, bone marrow, breast,
XX gastrointestinal tract, liver, lung, or urogenital; (b) immune
XX disorders e.g. Addison's disease, allergies, autoimmune haemolytic
XX anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
XX multiple sclerosis, rheumatoid arthritis and ulcerative colitis;
XX (c) cardiovascular disorders such as myocardial ischaemia; (d) wound
XX healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy;
XX and (f) infectious diseases such as viral, bacterial, fungal and
XX parasitic infections.
XX Note: The sequence data for this patent did not form part of the
XX printed specification, but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 73 AA;
SQ
Alignment Scores:
Pred. No.: 3,73e-10 Length: 73
Score: 19.00 Matches: 19
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.54% Indels: 0
DB: Gaps: 0
US-09-896-522-1 (1-1624) x ABB03475 (1-73)
QY 1521 CCACGTAGGAGTGTTCCTTCCCTTGTGCTTAAAGCAGAGACGAGCGG 1577
Db 11 ProlenAArgAspValPheProProCysLeuCysLeuLyGclYArgGluInlaArgArg 29
RESULT 15
ID ABLU2769 standard; Protein; 73 AA.
XX
XX ABLU2769;
AC
XX
XX 26-FEB-2003 (first entry)
DT
XX
XX Novel human musculoskeletal system antigen #389.
DE
XX

KM Musculoskeletal system antigen; cancer; metastasis;
 KM re-vascularisation; thrombosis; arteriosclerosis; mineral content;
 KM cardiovascular condition; wound; injury; burn; angiogenesis; ulcer;
 KM post-operative tissue repair; limb regeneration; neuronal growth;
 KM neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
 KM AIDS-related complex; chondrocyte growth; bone regeneration;
 KM periodontal regeneration; tissue transport; bone graft; skin aging;
 KM keratinocyte growth; hair loss; melanocyte growth; cell proliferation;
 KM cell growth; organ transplant; cell differentiation; body height;
 KM weight; hair colour; eye colour; skin; percentage of adipose tissue;
 KM pigmentation; cosmetic surgery; metabolism; biorythm; cardiac rhythm;
 KM depression; tendency for violence; pain; reproductive capability;
 KM hormone level; endocrine level; appetite; libido; memory; stress;
 KM storage capability; fat content; lipid content; protein content;
 KM carbohydrate content; vitamin content; cofactor content;
 KM nutritional component.
 XX Homo sapiens.
 XX OS
 PN US2002147140-A1.
 XX
 PD 10-OCT-2002.
 XX
 PF 17-JAN-2001; 2001US-0764877.
 XX
 PR 31-JAN-2000; 2000US-179065P.
 PR 04-FEB-2000; 2000US-180628P.
 PR 28-JUN-2000; 2000US-214886P.
 PR 07-JUL-2000; 2000US-216647P.
 PR 07-JUL-2000; 2000US-216880P.
 PR 11-JUL-2000; 2000US-217487P.
 PR 11-JUL-2000; 2000US-217496P.
 PR 14-JUL-2000; 2000US-218290P.
 PR 26-JUL-2000; 2000US-220963P.
 PR 14-AUG-2000; 2000US-224518P.
 PR 14-AUG-2000; 2000US-224519P.
 PR 14-AUG-2000; 2000US-225267P.
 PR 14-AUG-2000; 2000US-225268P.
 PR 14-AUG-2000; 2000US-225270P.
 PR 14-AUG-2000; 2000US-225477P.
 PR 14-AUG-2000; 2000US-225757P.
 PR 14-AUG-2000; 2000US-225758P.
 PR 22-AUG-2000; 2000US-226868P.
 PR 30-AUG-2000; 2000US-228924P.
 PR 01-SEP-2000; 2000US-229287P.
 PR 01-SEP-2000; 2000US-229343P.
 PR 01-SEP-2000; 2000US-229344P.
 PR 01-SEP-2000; 2000US-229345P.
 PR 05-SEP-2000; 2000US-228509P.
 PR 05-SEP-2000; 2000US-228513P.
 PR 08-SEP-2000; 2000US-231413P.
 PR 21-SEP-2000; 2000US-234274P.
 PR 21-SEP-2000; 2000US-234274P.
 PR 25-SEP-2000; 2000US-234997P.
 PR 27-SEP-2000; 2000US-235834P.
 PR 29-SEP-2000; 2000US-236327P.
 PR 29-SEP-2000; 2000US-236367P.
 PR 29-SEP-2000; 2000US-236368P.
 PR 29-SEP-2000; 2000US-236369P.
 PR 29-SEP-2000; 2000US-236370P.
 PR 02-OCT-2000; 2000US-236802P.
 PR 02-OCT-2000; 2000US-237037P.
 PR 02-OCT-2000; 2000US-237038P.
 PR 02-OCT-2000; 2000US-237039P.
 PR 13-OCT-2000; 2000US-237040P.
 PR 13-OCT-2000; 2000US-239335P.
 PR 20-OCT-2000; 2000US-240960P.
 PR 20-OCT-2000; 2000US-241785P.
 PR 20-OCT-2000; 2000US-241809P.
 PR 01-NOV-2000; 2000US-244617P.
 PR 17-NOV-2000; 2000US-249299P.
 PR 08-DEC-2000; 2000US-251856P.

PR 08-DEC-2000; 2000US-251868P.
 PR 08-DEC-2000; 2000US-251869P.
 XX
 PA (ROSE/) ROSEN C A.
 PA (RUBEN/) RUBEN S M.
 PA (BARA/) BARASH S C.
 XX
 PI Rosen CA, Ruben SM, Barash SC,
 XX WPI; 2003-128199/12.
 DR N-PSDB; ABX58045.
 XX
 PT Isolated nucleic acid molecules encoding musculoskeletal system
 associated polypeptides, useful for detecting disorders, e.g. cancer -
 XX
 PS Claim 11; SEQ ID NO 1422; 321pp; English.
 XX
 CC The invention describes an isolated nucleic acid molecule comprising a
 CC sequence encoding musculoskeletal system associated polypeptides useful
 CC for detecting disorders, e.g., cancer or cancer metastases, in animals
 CC or humans. The nucleic acid: stimulates re-vascularisation of ischaemic
 CC tissues associated with conditions such as thrombosis, arteriosclerosis,
 CC and other cardiovascular conditions; treats wounds due to injuries,
 CC burns, post-operative tissue repair, and ulcers; stimulates angiogenesis
 CC and limb regeneration; stimulates neuronal growth; can treat and prevent
 CC neuronal damage occurring in certain disorders or neurodegenerative
 CC conditions, such as, Alzheimer's disease, Parkinson's disease, and
 CC AIDS-related complex; stimulates chondrocyte growth, thus they can be
 CC used to enhance bone and periodontal regeneration and aid in tissue
 CC transports or bone grafts; prevents skin aging due to sunburn by
 CC stimulating keratinocyte growth; prevents hair loss, since PGP family
 CC members activate hair-forming cells and promotes melanocyte growth;
 CC stimulates growth and differentiation of hematopoietic cells and bone
 CC marrow cells when used in combination with other cytokines; maintains
 CC organs before transplantation or for supporting cell culture of primary
 CC tissues; induces tissue of mesodermal origin to differentiate in early
 CC embryos; increases or decreases the differentiation or proliferation of
 CC embryonic stem cells, besides, hematopoietic lineage, modulates
 CC mammalian characteristics, such as, body height, weight, hair colour, eye
 CC colour, skin, percentage of adipose tissue, pigmentation, size, and shape
 CC (e.g., cosmetic surgery); modulates mammalian metabolism; changes
 CC mammal's metal state or physical state by influencing biorythms;
 CC cardiac rhythms, depression, tendency for violence, tolerance for pain,
 CC reproductive capabilities, hormonal or endocrine levels, appetite,
 CC libido, memory, or stress; increases or decreases storage capabilities,
 CC fat content, lipid, protein, carbohydrate, vitamins, minerals, cofactors
 CC or other nutritional components. This is the amino acid sequence of a
 CC novel human musculoskeletal system antigen.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from the US patent office at
 CC ftp://seqdata.uspto.gov/sequence.html?DocID=20020147140.
 XX

Sequence 73 AA;

Alignment Scores:
 Pred. No.: 3.73e-10 Length: 73
 Score: 19.00 Matches: 19
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 3.54% Indels: 0
 DB: 24 Gaps: 0

US-09-896-522-1 (1-1624) x ABUI2769 (1-73)

Oy 1521 CCACTGAGGAGATTTCCTTCCCTTGTGCTTAAGGAGAGAGACGAGCGG 1577
 Db 11 ProleuhargAspValpheProfcysleuGlyArgGluAlaArgarg 29

Search completed: November 25, 2003, 07:53:13
 Job time : 78.4988 secs

This Page Blank (uspto)

Alignment Scores:

Pred. No.: 1,33e-72 Length: 201
Score: 1020.50 Matches: 200
Percent Similarity: 85.47% Conservative: 0
Best Local Similarity: 85.47% Mismatches: 1
Query Match: 34.40% Indels: 33
DB: 4 Gaps: 1

US-09-896-522-1 (1-1624) x Q96B10 (1-201)

QY 94 ATGGCTTCGGGCGGAGGCGGAAGACTGCGAGAGCCCGCGCGGAGCGGACCGTCCGAC 153
Db 1 MetalaeserlaglygluaspCysgluserProalaprogluialaapargProhis 20
QY 154 CAGGGCCCTTCCTGATAGGGGTGAGCGCGGCACTGCGACGGGAAAGTGCACCTGTGT 213
Db 21 GlnargProPheleuileglyvalserdlyglthralseserlylssethValCys 40
QY 214 GAGAAAGTCAATGAGTGTCTGGAGAGAAAGGAGGAGAAAGCGGAGCGGAGGAGTGTG 273
Db 41 GlnuylleuecgluLeuenglyGlnsmglnuvalGlnuInarglnInarglylval 60
QY 274 ATCTGAGCCAGGACAGGTTCTACAGAGTCTCTGACGAGAGCAAGAGCCGCTTG 333
Db 61 lleuserglInaapargpnetrylvalleuthrilaeglInuylalalyalaLeu 80
QY 334 AAAAGACATGTCATTTTGAACATCCAGATCCCTTGTATATGATTTGATGACAGACT 393
Db 81 LysgllyGlnlyrAanpneAspHisProAspAlaPheAspAenAspLeuMetHisargThr 100
QY 394 CTGAAGAACATCTGAGAGGCGAAAGCGGTGAGGCGGACCTTGTATTTTGTGACAC 453
Db 101 LeuylsasnilevalGlnuyllylThValGlnuvalProthrlyrAspneValThHis 120
QY 454 TCAAGTTACAGAGACCGAGCGGTGTCTACCTCGGACGAGTGTCTGTTGAGGAGATC 513
Db 121 SerargleupProgluInThrValVallyrProalAspValValleuphnegllyle 140
QY 514 TTGGTGTTCACACCGAGAGATCCGGGACATGTTCCACCTCGGCGCTTCTGTGACACC 573
Db 141 LeuvalPheylrserGlnuylleargAspMetPheHisleuArgleuPheValAspThr 160
QY 574 GACTCCGACGTCAGGCTGTCTCGAAGATTCTCGGAGCGTGGCGGAGGAGGACCTG 633
Db 161 AspserAspValAlaArgleuserArg-Arg----- 169
QY 634 GAGCAGATTCTGACGAGTACACCACTTCGTGAAGCCGCGCTTCAGAGAGTTCTGCTG 693
Db 169 ----- 169
QY 694 CCGCAAGAAGATGTCGGATGTGATCATCCACGAGAGTGCACATATGTTGCCATC 753
Db 170 --AspLysGlnuvalCysargCysAspHisProthrArgserGlnlyGlnlyCysHisG 189
QY 754 AACCTGATCGTGCAGACATCCAGACATTCATGATGG 791
Db 189 InProAspArgAlaAlaHisProGlnHisserGlnutr 201

RESULT 2
Q8C476 PRELIMINARY, PRT, 261 AA.
AC Q8C476:
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DE Uridine-cytidine kinase 2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;

RX MEDLINE=22354683; PubMed=12466851;
RA The PANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team,
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
DR EMBL: AK082837; BAC38646.1; -.
SQ SEQUENCE 261 AA; 29438 MW; 780AA3C298AA8153 CRC64;

Alignment Scores:

Pred. No.: 1.54e-65 Length: 261
Score: 932.00 Matches: 187
Percent Similarity: 79.04% Conservative: 28
Best Local Similarity: 68.75% Mismatches: 41
Query Match: 31.41% Indels: 16
DB: 11 Gaps: 4

US-09-896-522-1 (1-1624) x Q8C476 (1-261)

QY 115 GACTGGAGAGAGCCCGCGGAGGCGGACCGTCCGAC---CAGCGGCTTCTGATTA 171
Db 4 AspSerGlnGlnThrleuGlnasnHisGlnGlnProAsnGlyGlyGlnProPheleuile 23
QY 172 GGGGTGAGCGCGGACCTGCGACGCGGAAGTGCACCGTGTGAGAGATCAATGAGTTG 231
Db 24 GlyValSerGlyGlyThrAlaSerGlyLysSerSerValCysAlaAlaValleValGlnleu 43
QY 232 CTGGGACAGAACGAGGTGAAACAGCGGACGCGGAGGTGTATCTGTGACGAGACAG 291
Db 44 LeuGlyGlnasnGlnuvalAspTyrHisGlnlyGlnuValValleuserGlnAspser 63
QY 292 TTCTACAGGCTCTGAGCGGACGAGAGCGCAAGCGCTTGAAGACATGATTTT 351
Db 64 PheylrArgvalleuThrserGlnuyllylAlaValleuylsGlnPheAspne 83
QY 352 GACCATCCAGATGCTTTGATATGATTTGATGACAGACTCTGAAGACATCTGTGAG 411
Db 84 AspHisProAspAlaPheAspAenGlnleuilePheLysThrleuylsGlnleuThrGln 103
QY 412 GCGAAACGCGTGGAGGTCGACCTTATGATTTGTGACACATCAAGATTACAGAGACC 471
Db 104 GlyLysThrValGlnleProVallyrAspPheValSerHisSerArgLysGlnuThr 123
QY 472 AGCGTGTCTACCTCGGAGCGGTGTCTGTTGAGGAGATCTGTGTTCTACAGCCAG 531
Db 124 ValThrileylrProalAspValValleuphnegllyleuAlaPheylrserGln 143
QY 532 GAGATCCGGAGACATGTTCCACCTGCGCTTCTGTGAGACACGACTCCGACGTGAGCTG 591
Db 144 GlnuvalArgAspLeuPheGlnMetLysleuPheValAspThrAspAlaAspThrArgleu 163
QY 592 TCTCGAAGATTCTCGGAGACTG---CGCCGAGGAGGAGACCTGAGACAGATTCTGACG 648
Db 164 SerArgArgValleuArgAspLysSerGlnArgGlyArgAspLeuGlnGlnleuser 183
QY 649 CAGTACACACCTTGTGTAAGCGGCTTCGAGGATTCGCTGCGCAAGAAGATAT 708
Db 184 GlnlyrileThrPheVallyserProAlaPheGlnuyllylPheCysleuProthrlylslyr 203
QY 709 GCCGATGTGATCATCCACGAGAGTGCACATATGTTGCCATCACTGATCTGTGAG 768
Db 204 AlaAspValileleleProArglyAlaAspAenleuValAlaIleasnleuileValGln 223
QY 769 CACATCCAGACATTTGTAATGTGTGACATCTCAATATGCAACGAGAGGATTCATGGG 828
Db 224 HisileGlnAspLleleuasnGly-----Gly 232
QY 829 CGAGGTACAAAGCGGACCTTTCTGAGCAGGAGGACCAACCTGGAGTGTGACCTGTGG 888
Db 233 LeuSerLysArgGlnThr-----AsnGlyTyrPheAsnGlylyrThrProSerArg 249
QY 889 AACGCTCACATTGGAGTCCAGACAGACCCAC 924


```

Db      250 LysArgGlnAlaSerGluSerSerArgProHis 261
RESULT 3
Q8MRJ1 PRELIMINARY; PRT; 260 AA.
AC Q8MRJ1;
DT 01-OCT-2002 (TEMBLrel. 22, Created)
DT 01-OCT-2002 (TEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TEMBLrel. 23, Last annotation update)
DE L13909P.
GN C6364.
OS Drosophila melanogaster (fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Fafan D., Frise E.,
RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunco J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celinker S.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY119583; AM50237.1; -.
DR FlyBase; FBgn0039179; CG6364.
DR InterPro; IPR006083; PRK_URK.
DR InterPro; IPR00764; Uridine_kin.
DR Pfam; PF00485; PRK.1.
DR PRINTS; PRO0988; URIDINKINASE.
SQ SEQUENCE 260 AA; 29223 MW; 03CAA6DCA04A16B5 CRC64;

Alignment Scores:
Pred. No.: 1e-49 Length: 260
Score: 733.00 Matches: 142
Percent Similarity: 82.46% Conservative: 32
Best Local Similarity: 67.30% Mismatches: 35
Query Match: 24.71% Indels: 2
DB: Gaps: 2

US-09-896-522-1 (1-1624) x Q8MRJ1 (1-260)
QY 160 CCCTTCTGATAGGGGTGACGGCGGACTGCCAGCGGGAATCGACCTGTGTGAGAAG 219
DB 27 PripheleuilegilyalalagllyglYThralaserclYlysserThralValcyslysls 46
QY 220 ATCATGAGTGTGCGGAGCAAGAGGTGGAACAGCGGAGCGGGAAGGTGCATCTCG 279
DB 47 IlemecGlnleuGlyGlnAlaGlnMetCaspHisThrGlnArgGlnAlaValserIle 66
QY 280 AGCCAGGACAGGTCTTCAAGAGTCTGACCGGAGAGCAAGGCAAGGCTTGAAAGA 339
DB 67 SerGlnasperpherYrArGlnleuThrProAlaGlnlyslalysAlaGlnlyslY 86
QY 340 CAGTACATTTTGAACCATCCAGATGCTTTGATATATATTGATGACACAGACTCTGAAG 399
DB 87 LeupheasnphesphIsProAspAlaPheasnGlnGlnleuMetYrYrserThrLeuGln 106
QY 400 AACATCGTGGAGGCAAAACGATGAGGCGGACCGACCATGATTTTGTGACACACTCA--- 456
DB 107 AsnIleleuylsGlyslsYslsValGlnIleProserYrAspYrArGlnTrnAsnSerLeu 126
QY 457 AGGTTACAGAGACCAACGATGATCTACCTGCGGACCTGTGTTGAGGCGATCTTG 516
DB 127 AspPheGlnuasnValleuValIleYrProAlaAspAlaValleuPheGlnGlyIleLeu 146
QY 517 GTGTTCTACAGCCAGAGATCCGGGACATGTTCCACTCGGCGCTCTTCTGTGACACCGAC 576
DB 147 ValPheYrIlePheProYrslleArgGlnleuPheHleMetYrleuPheValaAspTrnAsp 166
QY 577 TCGACGTCAGGCTGTCTCGAAGAGTTCTCCGGGACGTG---CGCCGAGGAGGACCTG 633

```

```

Db      167 ProAspThrArgLeuAlaArgArgValProArgAspIlaAsnGlnArgGlyArgAspLeu 186
QY 634 GAGCAGATTCTGACGACGATPACACCACTTGTGTGAACCGCGCTTCGAGAGATTCTGCTG 693
DB 187 AspAlaValleuThrGlnTrnMetThrPheValYsProAlaPheGlnGlyIlePheCysSer 206
QY 694 CGGACAAAGAGATGAGCCCATGATCATCCACGAGAGTGAGCAATATGTGTTGCCATC 753
DB 207 ProThrLysAlaPheAlaAspValIleIleProArgGlyAlaAspAsnThrValaIaIle 226
QY 754 AACCTGATCGTACGACATCCAGCATTCCTG 786
DB 227 AspLeuIleValaHisHisIleGlyIleLeu 237

RESULT 4
Q9FKS0 PRELIMINARY; PRT; 486 AA.
AC Q9FKS0;
DT 01-MAR-2001 (TEMBLrel. 16, Created)
DT 01-MAR-2001 (TEMBLrel. 16, Last sequence update)
DT 01-MAR-2003 (TEMBLrel. 23, Last annotation update)
DE Uridine kinase-like protein.
GN AT5G40870.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Columbia;
RX MEDLINE=98344145; PubMed=9679202;
RA Kaneko T., Kotani H., Nakamura Y., Sato S., Asamizu E., Miyajima N.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. V. Sequence
RT features of the regions of 1,381,565 bp covered by twenty one
RT physically assigned pl and TAC clones.";
RL DNA Res. 5:131-145 (1998).
RN [2]
RP SEQUENCE FROM N.A.
RA Yamada K., Chan M.M., Chang C.H., Dale J.M., Hsuan V.W., Lee J.M.,
RA Onodera C.S., Quach H.L., Tang C.C., Toriumi M., Wong C., Wu H.C.,
RA Yu G., Yuan S., Chen H., Cheuk R., Jones T., Kim C.J., Nguyen M.,
RA Palm C.J., Shim P., Southwick A., Tripp M.G., Wu T., Davis R.W.,
RA Ecker J.R., Theologis A.;
RT "Arabidopsis Open Reading Frame (ORF) Clones.";
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB011477; BAB1349.1; -.
DR EMBL: BT002336; AM86169.1; -.
DR HSSP; Q26998; 1BD3.
DR InterPro; IPR006083; PRK_URK.
DR InterPro; IPR000836; PRTtransferase.
DR InterPro; IPR00764; Uridine_kin.
DR Pfam; PF00156; Priboylcran.1.
DR Pfam; PF00485; PRK.1.
DR PRINTS; PRO0988; URIDINKINASE.
DR TIGRPFAMs; TIGR00235; udk.1.
KW Kinase; Transferase.
SQ SEQUENCE 486 AA; 54430 MW; 50DA1CE89346FB54 CRC64;

Alignment Scores:
Pred. No.: 7.42e-37 Length: 486
Score: 572.00 Matches: 123
Percent Similarity: 61.38% Conservative: 55
Best Local Similarity: 42.41% Mismatches: 76
Query Match: 19.28% Indels: 36
DB: Gaps: 8

US-09-896-522-1 (1-1624) x Q9FKS0 (1-486)
QY 7 TCGCCTCCGACCTCGGCGGTGGCGGCGCGCGGCGCGGCGGCGGCGGCGGCGGCGGCGG 66

```

```

Db      32 SerProProAnSerSerVal-----38
Qy      67 ACCGGATGCGGGAGCGGAGCCGAGATGCTTGGCGGAGCGAAGCTGCGAGAC 126
Db      39 -----ValSerSerLeuArgSerAlaValSerSerSerProSerSerSerAapPro 56
Qy      127 CCGCGCGGAGGCGGACCGTCCGACGAGCGCTTCCGATGAGGGGTAGGGGGGCGC 186
Db      57 GluAlaPro-----LysGlnProPheIleIleGlyValSerGly 70
Qy      187 ACTGCGACGGGAATGCGACCGTGTGTGAGAGATCATGAGATTGCTGGACAGAACGAG 246
Db      71 ThrAlaSerGlyLysThrThrValCysAspMetIleIleGlnGlnLeu-----86
Qy      247 GTGAAACAGCGGACGGAGAGTGTCTCTGAGCCAGGAGTTTACAAAGTCTG 306
Db      87 -----HisAspHisArgValAlaValLeuValAsnGlnAspSerPheTyArgGlyLeu 103
Qy      307 ACGGAGAGCGAAGGCGGACCGCTTGAAGAAGACATATTTTGAACATCCAGAGTCC 366
Db      104 ThierGlnGlnLeuGlnArgVal-----GlnGlnTyAsnPheAspHisProAspAla 121
Qy      367 TTGATATATGATTGATGACAGACACTCTGAAGAACATCTGAGGCGAAGCGTGAG 426
Db      122 PheAspThrGlnGlnLeuLeuHisGlyAlaGlnThrLeuLysSerGlyGlnProTyGln 141
Qy      427 GTGCGGACCTTATGATTGTTGACACATCAAGTTTACAGAGACC--ACGGTGTCTAC 483
Db      142 ValProIleTyAspPheLysThrHisGlnArgSerAspThrPheArgGlnValAsn 161
Qy      484 CCGCGGACGAGTGTCTGTTGAGGACATCTTGGTGTCTACGACGAGAGATCCGGGAC 543
Db      162 AlaSerAspValIleIleLeuGlnGlyLeuValPheHisAspSerArgValAlaArgAsn 181
Qy      544 ATGTTCACCTGCGCTTCTGTGACACCGACTCCGAGCTGTCTCGAAGATT 603
Db      182 LeuMetAspMetLysIlePheValAspThrAspAlaAspValAlaArgLeuAlaArgGlyLe 201
Qy      604 CTCCGGGAC--GTGCGCGGAGGAGGACCTGACGACGATTTGACGCGCATACCAACC 660
Db      202 ArgArgAspThrValGlnArgGlyArgAspValAsnSerValLeuGlnGlnTyAlaLys 221
Qy      661 TTGCGAAGCGGCGCTTGTGAGAGATGCTGCGCGACCAAGATATGCGGATGTATC 720
Db      222 PheValLysProAlaPheAspAspPheValLeuProSerLysTyAlaAspValIle 241
Qy      721 ATCCACGAGGAGTGACCAATATGTTGCCATCAACTGATCTGACAGACATCCAGGAC 780
Db      242 IleProArgGlyGlyAspAsnHisValAlaValAspLeuIleThrGlnHisIleHisThr 261
Qy      781 ATTGTG--AATGTGACATCTGCAAAATGGACCGGAGAGGTCCATGGCGGAGCTAC 837
Db      262 LysLeuGlnGlnHisAspLeuCysLysIleTy-----ProAsnValTyArgValIle 278
Qy      838 AAGCGACCTTTCTGAGCCAGGAGCAC 867
Db      279 GlnSerThrPheGlnIleArgGlyMetHis 288

RESULT 5
Q9LTY6 PRELIMINARY; PRT; 465 AA.
AC Q9LTY6;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Uridine kinase-like protein.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eustosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_taxid=3702;
RN
RP SEQUENCE FROM N.A.

```

```

RC STRAIN=Columbia;
RA Sato S., Nakamura Y., Kaneko T., Kato T., Asamizu E., Tabata S.;
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Columbia;
RX MEDLINE=20277480; PubMed=10819329;
RA Nakamura Y.;
RT "Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence
RT features of the regions of 4,504,864 bp covered by sixty P1 and TAC
RT clones."
RL DNA Res. 7:131-135(2000).
DR EMBL: AB024028; BAA95720.1; -
DR HSP: Q26998; 1BD3
DR InterPro: IPR006083; PRK URK.
DR InterPro: IPR000764; Uridine_kin.
DR Pfam: PF00485; PRK.1.
DR PRINTS: PR00986; URIDINKINASE.
DR TIGRFAMs: TIGR00235; udk; 1.
KW Kinase.
SO SEQUENCE 465 AA; 52125 MW; 4CD82DB9059E7168 CRC64;

Alignment Scores:
Pred. No.: 1,52e-36 Length: 465
Score: 568.00 Matches: 115
Percent Similarity: 66.94% Conservative: 51
Best Local Similarity: 46.37% Mismatches: 68
Query Match: 19.14% Indels: 14
DB: Gaps: 5

US-09-896-522-1 (1-1624) x Q9LTY6 (1-465)
Qy 130 GCCCGGAGGCGGACCGTCCGACCGGCGCTTCTGATAGGGTGAACGGCGGCGACT 189
Db 19 SerProSerAlaProAlaProAlaProLeuLysGlnProPheValIleGlyAlaGlyGlyThr 38
Qy 190 GCCAGCGGAGTGCAGACCGTGTGTGAGAGATCATGAGATTGCTGGACAGAACGAGTG 249
Db 39 AlaSerGlyLysThrThrValCysAspMetIleMetSerGlnLeu-----53
Qy 250 GAACAGCGGCGGAGGAGTGTCTATCTGAGCCAGAGAGTCTTACAAAGTCTGACG 309
Db 54 -----HisAspLysArgValValLeuValAsnGlnAspSerPheTyHisSerLeuThr 71
Qy 310 GCAGACGAGAGGCCAAGCGCTTGAAGAAGACATCAATTTGACATCCATGAGTCTT 369
Db 72 LysGlnLysLeuAsnLysVal-----HisGlnTyAsnPheAspHisProAspAlaPhe 89
Qy 370 GATATGATATTTGATGACAGACTGTAAGAACATCATCTGGAGGCGCAAAACGGTGAGTG 429
Db 90 AsnThrGlnValLeuLeuSerCysMetGlnLysLeuArgSerGlyGlnProValAsnIle 109
Qy 430 CCGACCTATGATTTTGTGACACATCAAGTTTACAGAACACCGTGTCTTACCTCGG 489
Db 110 ProSerTyAspPheLysIleHisGlnSerIleGlnLysSerSerProValAsnProGly 129
Qy 490 GAGGTGTTCTGTTTGAAGGCACTCTGTTGTTTACAGCCAGAGATCCGGGACATGTC 549
Db 130 AspValIleIleLeuGlnGlyIleLeuValLeuAsnAspProArgValAlaArgAspLeuMet 149
Qy 550 CACCTGCGCTCTCTGTGACACCGACTCGAGCTGAGCTGTCTGAGAGAGTTCTCCG 609
Db 150 AsnMetLysIlePheValAspThrAspAlaAspValAlaArgLeuSerArgGlyIleGlnArg 169
Qy 610 GAC--GTGCGCGGAGGAGGAGACTGAGACAGATTCTGACGAGTACACCACTTCTGTG 666
Db 170 AspThrValGlnArgGlyArgAsnIleGlnAsnValLeuGlnGlnTyThrLysPheVal 189
Qy 667 AACCGCGCTTGAAGAGTTCCTGCGCGGACAAAGATGATGATGATGATGATGATGAT 726
Db 190 LysProSerPheAspGlnTyArgIleGlnProSerMetLysTyAlaAspIleIleLeuPro 209
Qy 727 CGAGAGTGAACATATGTTGCTCATCAACTGATCTGTCAGACATCCAG--GACATT 783

```

```

Db      210 ArgGlyValSerPheAspValAlaIleAspLeuIleValGlnHisIleArgThrLysLeu 229
Qy      784 CTGAATGTCACATCTCCAAATGACGAGAGGCTCCAAATGGCGGAGCTTACAAACGG 843
Db      230 CysGlnHisAsnLeuCysLysIleTyr-----SerAsnIlePheIleIleSerSer 246
Qy      844 ACCTTCTGAGCCGAGGAGCCAC 867
Db      247 ThrPheGlnIleLysGlyMetHis 254

RESULT 6
Q8N524
ID      Q8N524      PRELIMINARY;      PRT;      419 AA.
AC      Q8N524;
DT      01-OCT-2002 (TrEMBLrel. 22, Created)
DT      01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DE      01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE      Similar to uridine kinase-like 1.
OS      Homo sapiens (Human) .
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX      NCBI_TaxID=9606;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      TISSUE=Placenta;
RA      Strausberg R.;
RL      Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR      EMBL; BC033078; AAH33078.1; -
DR      InterPro; IPR006082; PRK.
DR      InterPro; IPR006083; PRK URK.
DR      InterPro; IPR000764; Uridine_kin.
DR      Pfam; PF00485; PRK; 1.
DR      PRINTS; PR00478; PRIBDKINASE.
DR      PRINTS; PR00988; URIDINKINASE.
DR      TrRfam; TRFR00235; udk; 1.
KM      Kinase.
SQ      SEQUENCE 419 AA; 45972 MW; 1D1F1E75305AD5 CRC64;

Alignment Scores:
Pred. No.:      2,81e-36      Length:      419
Score:      564.50      Matches:      148
Percent Similarity:      46.70%      Conservative:      57
Best Local Similarity:      33.71%      Mismatches:      125
Query Match:      19.03%      Indels:      109
DB:      Gaps:      13

US-09-896-522-1 (1-1624) x Q8N524 (1-419)
Qy      7 TCGCCTCCGACCTCGGCGCTGGCGGCGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGG 66
Db      16 SerProThrAla-----ArgAspThrProGlyArg----- 26
Qy      67 ACCGATCGCGGAGGAGCGGAGCGGAGATGATGCTTCGCGGAGGCGAAGAC----- 117
Db      27 -----GlnAlaGluLysSerGlnThrAlaCysGlnAspArgSerAsn 40
Qy      117 ----- 117
Db      41 AlaGlnSerLeuAspArgLeuLeuProValGlyThrGlyArgSerProArgLysArg 60
Qy      118 -----TCGAGAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 135
Db      61 ThrThrSerGlnCysLysSerGlnProProLeuLeuArgThrSerLysArgThrIleTyr 80
Qy      136 GAGCGCGACCGCTCG-----CACGAGCGCGCGCTCTCTG 168
Db      81 ThrAlaGlyArgProProTyrArgAsnGlnHisGlyThrGlnSerLysGlnAlaPheAla 100
Qy      169 ATGAGGCTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 228
Db      101 IleGlyLeuGlyLysSerAlaSerGlyLysThrThrValAlaArgMetIleIleGlu 120

```

```

Qy      229 TTCGTGGACAGAACGAGGTGGAAACGCGACGCGAGAGGTGTGATCTGACCGACGAC 288
Db      121 AlaLeu-----AspValProTyr-----ValValLeuLeuSerMetAsp 133
Qy      289 AGGTTTCAAGGCTCTTGAACGCGACGACGAGAGCGCAAGGCTTGAAGACATCAAT 348
Db      134 SerPheTyrLysValLeuThrGlnGlnGlnGlnGlnAlaIleAsnAsnPheAsn 153
Qy      349 TTTGACCATTCGAGAGGCTTGTGATATGATTTGATGACACATCAAGATTACAGAG 408
Db      154 PheAspHisProAspAlaPheAspPheAspLeuIleIleSerThrLeuLysLysLys 173
Qy      409 GAGCGCAAAACGCTGAGGTGCGGACCTGATGATTTGTGACACATCAAGATTACAGAG 468
Db      174 GlnGlyLysSerValLysValProIleTyrAspPheThrThrHisSerArgLysAsp 193
Qy      469 ACCAGCGGTCTTACCGCTTCGCGAGCGGTGCTTTGAGGCGATCTTGTTCTACAGC 528
Db      194 TrpLysThrLeuTyrGlyAlaAsnValIleIlePheGlnGlyLysMetAlaPheAlaAsp 213
Qy      529 CAGGAATCCGCGACATGTTCCATGCGCGCTTGTGTCGACACCGACCTCCGACGACAG 588
Db      214 LysThrLeuLeuGlnLeuLeuAspMetLysIlePheValAspThrAspSerIleArg 233
Qy      589 CTGTCTGAAAGATCTCCGCGGACGTG---CGCGAGGAGGAGGACCTGGACGATTTCTG 645
Db      234 LeuValAlaArgThrLeuArgThrAspIleSerGlnArgGlyAlaAspIleGlnGlyValIle 253
Qy      646 ACAGCATACACACCTTCGTGAAACCGGCGCTTCGAGAGATTCGCTCCGACAAAGAG 705
Db      254 LysGlnTyrAsnLysPheValLysProSerPheAspGlnTyrIleGlnProThrMetArg 273
Qy      706 TATCGCGATGTATCATCTCCACGAGAGTGACAAATATGTTGCCATCAACCTGATCTG 765
Db      274 LeuAlaAspIleValValProArgGlySerGlyAsnThrValAlaIleAspLeuIleVal 293
Qy      766 CAGCATCATCAGACATCTGTAATGTGATCATCTGCAATGGACCGAGAGGCTCCAAAT 825
Db      294 GlnHisValHisSerGlnLeu-----GlnGlnGlyCysAla 305
Qy      826 GGGCGGAGGTACAAAGCGACCTTTTCTGAGCCA---GGGAGCACCTTGAGATGTGAGC 882
Db      306 GlyLeuGlyThrProValProProAlaIaProAspAlaGlnArgProGlnGlnHisAla 325
Qy      883 TCTGCGAAACGCTCATTTGGATGGATCCAGACGACGCCCATGAGGCGCTCCGAGCT 942
Db      326 AlaGlyThrGlyHisAlaHisHisGlnValArgAlaHisIleuGlyThrGlyArgPro 345
Qy      943 CAGGCGAGGTCTCCGCGCGGACATGTGTGTCAGGACGTGAGGCTGGGACGCCACCCA 1002
Db      346 ArgAlaArgAlaProAlaGln----- 352
Qy      1003 CACCCATGCTCTCTCTCGCGCGACCCCGAGGAGTGTAGCAGCGAGCC----- 1053
Db      353 -----LeuLeuAlaProGlnGlnGlnGlyAspGlnSerArgArgValHis 367
Qy      1054 TTCCTCATCTGAGAGTGGAAACTCAGATGTGTCACTCAAGATCACTTGCTGGGACACTG 1113
Db      368 LeuLeuLeuGlnIleThrAspAlaAlaAlaHisArgAlaArgAlaLeuLeuProAlaLeu 387
Qy      1114 ACAAGGCTTCGAGGTTTTCAGCCACTTAAGCTGTCGGGTTTAAGATCCCTCT 1170
Db      388 SerGly-----LeuArgArgThrAspProAla 396

RESULT 7
Q9LK34
ID      Q9LK34      PRELIMINARY;      PRT;      483 AA.
AC      Q9LK34;
DT      01-OCT-2000 (TrEMBLrel. 15, Created)
DT      01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE      01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE      Uridine kinase-like protein.
OS      Arabidopsis thaliana (Mouse-ear cress) .

```

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCB1_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Columbia;
 RA Kaneko T., Kato T., Sato S., Nakamura Y., Asamizu E., Tabata S.;
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Columbia;
 RA MEDLINE=20363099; PubMed=10907853;
 RT "Structural analysis of Arabidopsis thaliana chromosome 3. II.
 RT Sequence features of the regions of 4,251,695 bp covered by ninety PL,
 RT TAC and BAC clones";
 RL DNA Ref. 7:217-221(2000).
 DR EMBL; AP000381; BAB02114.1; -.
 DR HSSP; Q26998; 1BD3.
 DR InterPro; IPR006083; PRK URK.
 DR InterPro; IPR000764; Uridine_kin.
 DR Pfam; PF00485; PRK; 1.
 DR PRINTS; PRO0988; URIDINKINASE.
 DR TIGRFAMs; TIGR00235; udk; 1.
 KW Kinase.
 SQ SEQUENCE 483 AA; 54210 MW; 2862F7AED187020F CRC64;
 Alignment Scores:
 Pred. No.: 3.85e-36 Length: 483
 Score: 563.00 Matches: 117
 Percent Similarity: 67.08% Conservative: 46
 Best Local Similarity: 48.15% Mismatches: 16
 Query Match: 18.98% Indels: 16
 DB: 10 Gaps: 7
 US-09-896-522-1 (1-1624) x Q9LK34 (1-483)
 QY 148 CCGGACCCGCGCCCTTCTGATAGGCGTGAAGCGGCGGACCTGCGGAGGAGTGCACC 207
 Db 59 Protheglin---Protheglinlethgylvalthrglygylthrralasegilylthrr 77
 QY 208 GTGTGTGAAGAGTCAATGAGTGTGCGGACGAGGAGTGAACAGCGGCGGCGGAG 267
 Db 78 ValCysaspmetleileglnleu-----Hisasphlary 90
 QY 268 GTGTGATCTGCGGCGGAGGAGTGTGACAGGCTGAGCGGAGGAGGAGGAGGAGGAG 327
 Db 91 Ilevallleuvalasnghlaspserphetyrarglyleuthrsergluylleu 110
 QY 328 GCGCTTGAAGGAGGAGTCAATGAGTGTGAGGAGTGTGAGTGTGAGTGTGAGTGTG 387
 Db 111 Val-----GlnghuylrAsnPhelaspHisProasphalrPhelaspThnglnleu 128
 QY 388 AGGACTGTGAAGAGGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 447
 Db 129 HisCysvalleuvalleuylserglglnprrglnleproilleyrAspPhelys 148
 QY 448 ACAGACTGAAGGAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 504
 Db 149 ThrhisglnarglylvalaspalrPhelaspThnglnleuvalasnghlaspValleu 168
 QY 505 GAGGAGGAGTGTGATCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 564
 Db 169 GluylleuvalleuvalrpheniaspserargylargharspMetasMetlyllephe 188
 QY 565 GTGACACCGGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 621
 Db 189 ValaspThrAspAlaAspValargleuvalarglileargAspArpThValgluarg 208
 QY 622 GGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 681
 Db 209 GlyArgAspValAspSerValleuGlnGlnThrValalysPheValLysPheAlaPheAsp 228

QY 682 GAGTTCGCTGCGGAGCAAAAGATATGCCGATGATCATCCACGAGGAGTGCACAT 741
 Db 229 AspPheValleuProserlyslrYrAlaAspValillelrohrnglyglAspAsn 248
 QY 742 ATGTTGCCATCAACTGATGTCGAGACACATCCAGGACATTTCG---AATGTGACATC 798
 Db 249 HisValAlaValAspPheulleValglnHisleHisThrlyleuGlyglnHisAspLeu 268
 QY 799 TGCAGATGCGACGAGGAGGAGTGCATGCGGAGGAGTGCACAGGAGGAGGAGGAG 858
 Db 269 Cyslylellyr-----ProasPhalPheValillelGlnThrPheGlnllearg 285
 QY 859 GGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 867
 Db 286 GlyMetHis 288
 RESULT 8
 Q19583 ID Q19583 PRELIMINARY; PRT; 515 AA.
 AC Q19583;
 DT 01-NOV-1996 (TREMURel. 01, Created)
 DT 01-NOV-1996 (TREMURel. 01, Last sequence update)
 DT 01-MAR-2003 (TREMURel. 23, Last annotation update)
 DE F19B6.1a protein.
 GN F19B6.1 OR F19B6.1A.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
 OC Rhabditidae; Peloderrinae; Caenorhabditis.
 OX NCB1_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Thomas K.;
 RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=99069613; PubMed=9851916;
 RA none;
 RT "Genome sequence of the nematode C. elegans: A platform for
 RT investigating biology";
 RL Science 282:2012-2018(1998).
 DR EMBL; Z69635; CA93459.1; -.
 DR HSSP; Q26998; 1BD3.
 DR WormPep; F19B6.1a; CE05666.
 DR InterPro; IPR006082; PRK.
 DR InterPro; IPR006083; PRK URK.
 DR InterPro; IPR000764; Uridine_kin.
 DR Pfam; PF00485; PRK; 1.
 DR PRINTS; PRO0478; PHRIBLKINASE.
 DR PRINTS; PRO0988; URIDINKINASE.
 DR TIGRFAMs; TIGR00235; udk; 1.
 SQ SEQUENCE 515 AA; 58452 MW; 9584D947A2D3B5C5 CRC64;
 Alignment Scores:
 Pred. No.: 5.15e-36 Length: 515
 Score: 561.50 Matches: 125
 Percent Similarity: 63.10% Conservative: 46
 Best Local Similarity: 46.13% Mismatches: 62
 Query Match: 18.92% Indels: 38
 DB: 5 Gaps: 6
 US-09-896-522-1 (1-1624) x Q19583 (1-515)
 QY 12 TCGGACCTGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 68
 Db 14 SerProargAlaAlaClyCysargThrArgArgArgThMetserlyglYrAla 33
 QY 69 CCGATGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 128
 Db 34 -----HisHisleuLeuThr 39
 QY 129 GGGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 153
 Db 153 -----HisHisleuLeuThr 39

Db	40	lystHnGlyLysLysLysLeIytrHrthlySgLYaNgPrO-ProTPrTyAspRlySgLY	55
QY	154	-----CAGCGCCCTTCTGTATAGGGGTGAAGCGCGGCACTGCCAGCGGAAGTCGAC	206
Db	59	serLeuYshISrProPheValIleGlyalCySgLYgLYserAlaserGlyystrH	79
QY	207	CSTGTGTGAAGATPOTAGAGTGTCTGGGACAGAAACAGAGTGAACAGCGGACGCGAA	266
Db	79	rValIaGlyLysIleValIGlyuNgLY-----IleProTr	92
QY	267	GGTGGCATCCCGAGCGAGCAAGTGTTCACAGAGTCCGTGAGCGAGACAGAAAGCGCA	326
Db	92	rValHrthIleLeuSerMetAspserPheTytrYsValIleThrProGluGluIleYsAl	112
QY	327	GGCCTTGAAAGGACATACATATTTTGACCATCCAGATGCCCTTGATATATGATTTGACA	386
Db	112	aAlaHsGlyserArgTyAsnRheAspGlyProaAlaIaPheAspRheAspLeuY	132
QY	387	CAGGACTCTGAAGACATGTGGAGGGCAAAACSGTGAAGCTGCCACCTATGATTTGT	446
Db	132	rGluValLeuYsArgLeuAdrgIuGlyuSerValAspValProValTyAspPheAs	152
QY	447	GACACATCAAGTTATCCAGAGCCAGCGGCTGTCACCTGGAGAGTGGTCTGTGTA	506
Db	152	nThrIshSerIaGAspRrAsnserYusMetCtyGtLYaIaAspAlaLeuIlePegI	172
QY	507	GGGCACTGTGTGTCTTACAGCCAGAGAGATCCCGGACATGTTCCACTCGACCTTCGT	566
Db	172	uGlyIleLeuAlaRheHiaAspGluArgIleYsAlaPheLeuMetAspMetLYsValIaPheVa	192
QY	567	GGACACCCGATCCGACAGTGAAGCTGTCTGGAAGAGTTCTCCGGAGCTGGCG--CGAGG	623
Db	192	IAspThrAspGlyAspLeuAdrgLeuAlaIaArgIleValaIaGAspValIthrAspArgI	212
QY	624	GAGGAGCCGCGAGCAATCTGACCGAGTAAACACCACTTCGTGAACCGGCTTCAGAGA	683
Db	212	YArgAspIleAspGlyIleMetGluGlyIleThrIlePheValIleYsProIaIaPheAspLy	232
QY	684	GTTTCGCTCGCGACAAAGATGATCCGATGATGATCATCTCCAGAGAGTGAACAATAT	743
Db	232	srYrIleAlaProCySmetAspSerAlaAspLeuIleValProArgGlyGlyGluAsnAs	252
QY	744	GGTGGCATCAACCTGATCGTGACGACATC	774
Db	252	rValaIaIaIeAspMetIleValIGlnsVal	262
RESULT 9			
QY	93	PRELIMINARY; PRT; 555 AA.	
AC	09U317;		
DT	01-MAY-2000 (TREMBLrel. 13, Created)		
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)		
DT	01-MAR-2003 (TREMBLrel. 23, Last annotation update)		
DE	F19B6.1b Protein.		
GN	F19B6.1 OR F19B6.1B.		
OS	Caenorhabditis elegans.		
OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;		
OC	Rhabditidae; Peloderinae; Caenorhabditis.		
OX	NCBI_TaxID=6239;		
RN	SEQUENCE FROM N.A.		
RP	Thomas K.;		
RL	Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=99069613; PubMed=9851916;		
RA	none;		
RT	"Genome sequence of the nematode C.elegans: A platform for		
RL	investigating biology.";		
RL	Science 282:2012-2018 (1998).		
DR	EMBL; Z69635; CAA93462.1; -.		
DR	HSSP; Q26998; 1BD3.		
DR	WormPep; F19B6.1b; CE20700.		

DR	InterPro: IPR006082; PRK.	
DR	InterPro: IPR006083; PRK URK.	
DR	InterPro: IPR00764; uridine_kin.	
DR	Pfam: PF00485; PRK; 1.	
DR	PRINTS: PRO0478; PHRIKINASE.	
DR	PRINTS: PRO0988; URIDINKINASE.	
DR	TIGRFAMs: TIGR00235; udk; 1.	
DR	SEQUENCE 555 AA; 62673 MW; D0786FAB98B8CF98 CRC64;	
Alignment Scores:		
Pred. No.:	5,266-36	Length: 555
Percent Similarity:	561.50	Matches: 125
Best Local Similarity:	63.10%	Conservative: 46
Query Match:	46.13%	Mismatches: 62
DB:	18.92%	Indels: 38
	5	Gaps: 6
US-09-896-522-1 (1-1624) x Q9U317 (1-555)		
QY	12 TCQCACTCCGCGCTGGG---GGGCGCGCCGGGCGGAGGAGGCGGCGCGGAGC	68
DB	54 SerProArgAlaAlaGlyCysArgThrArgArgThrMetSerGlyArgAlaGlu	73
QY	69 CGAGTCCGCGGAGGCGGAGGCCGGAATGCTTCGGCGGAGCGGAAGATGCCAGAGCC	128
DB	74 -----HisHisLeuThrThr	79
QY	129 GCGCGCGGA-----GGCGACCGTCCGAC-----	153
DB	80 LysThrGlyLysLysIleThrLysGlyArgPro-ProTyrLysLysGlyLys	99
QY	154 -----CAGCGCGCCCTTCTGATAGGGGTGAGCGCGGCACCTGCCAGCGGAGTGCAC	206
DB	99 sserLeuLysHisProPheValIleGlyValCysGlyGlySerAlaSerGlyLysThrTh	119
QY	207 CGTGTGTGAAGATCATGGAATTGCTGGGAGCAAGACGAGTGAACAGCGGACGCGAA	266
DB	119 rValIaGluLysIleValaGluArgLeuGly-----IleProtr	132
QY	267 GGTGTGATCCCGAGCCAGGACAGGTTCTCAAGGCTCGAGCGGACAGAGAGCCAA	326
DB	132 PalThrIleLeuSerMetAspSerPheThrLysValLeuThrProGluGluIleLysAl	152
QY	327 GGCCTTGAAGAGCAGTCAATTTTGAACCATCCAGATGCGCTTGATTAAGATTGATGA	386
DB	152 AlaHisGluSerAspGlyTrpAsnPheAspGlyProAsnAlaPheAspPheAspLeuLeu	172
QY	387 CAGGACTGGAAGAACATGTTGAGGAGGCAAAACGTTGAGAGTGCACCTATGATTTGT	446
DB	172 rGluValLeuLysArgLeuArgGluGlyLysSerValAspValProValTyrAspPheAs	192
QY	447 GACAACTCAAGTTACCAAGAGACCAACGAGTGGTCACTCGGGAGAGTGTCTGTGTA	506
DB	192 nThrHisSerArgAspProAsnSerLysMetMetLysGlyAlaAspValLeuIlePheGlu	212
QY	507 GGGCACTTGGTGTTCTACAGCCAGAGAAATCCGGGACATGTTCCACTCGGCTTTCGT	566
DB	212 uGlyIleLeuAlaPheHisAspGluArgIleLysAsnLeuMetAspMetLysAlaIheVa	232
QY	567 GGACACCGACTCCAGCGTCAAGGCTGTTCGAAGAGTTCTCCGGAGCGTGCGC---CGAGG	623
DB	232 LAspThrAspGlyAspLeuArgLeuAlaArgTrgIleValArgAspValThrAspArgGlu	252
QY	624 GAGGACCTGAGAGCAGATTTCAGCGCAGTACACCACTTCGTGAAGCCGGCTTCGAGGA	683
DB	252 YArgAspIleAspGlyIleMetGluGlnIleThrPheThrIleValLysProAlaPheAsp	272
QY	684 GTTTCGCTCGCGCAAAAGATAGTACGCAATGATGCATCCACGAGAGTGGACATAT	743
DB	272 sTyrIleAlaProCysMetAspSerAlaAspMetIleValProArgGlyGlyLysAsnAs	292
QY	744 GATTGCATCAACCTGATCGTGACACATC	774

Db 292 pValAla1leHisPheValGlnAsnVal 302

RESULT 10

Q9V810 PRELIMINARY; PRT; 614 AA.

AC Q9V810, Q9V811; (TREMBlrel. 13, Created)

DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)

DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)

DE CG4798 protein.

GN CG4798

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Peerygota;

OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI TaxID=7227;

RN [1]

RP SEQUENCE FROM N.A. (LONG AND SHORT ISOFORMS).

RC STRAIN=BERKELEY;

RX MEDLINE=20196006; PubMed=10731132;

RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,

RA Amanatides P.G., Scherer S.E., Li P.W., Hoeklin R.A., Galle R.F.,

RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,

RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,

RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,

RA Ballew R.M., Baus A., Baxendale J., Bayraktaroglu L., Beasley E.M.,

RA Boekova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,

RA Butler J.K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,

RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,

RA de Pablos R., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,

RA Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,

RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrieres S., Fleischmann W.,

RA Fodde C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,

RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan X., Harris M.,

RA Harris N.L., Harvey D., Heilmann T.J., Hernandez J.R., Houck J.,

RA Hootin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,

RA Jaitai M., Kalush F., Karpen G.H., Ke Z., Kemison J.A., Ketchum K.A.,

RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

RA Laoko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,

RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,

RA Mikhlov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,

RA Mount S.M., Moy M., Murphy B., Murphy D.M., Nelson D.L.,

RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Paclob J.M.,

RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,

RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,

RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,

RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,

RA Svrtkars R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,

RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weisenbach J.,

RA Williams S.W., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,

RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,

RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,

RA Gbbs R.A., Myers E.W., Rubin G.M., Venter J.C.;

RT "The genome sequence of Drosophila melanogaster.";

RL Science 287:2185-2195(2000).

CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: A LONG FORM (SHOWN HERE) AND A

CC SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.

DR EMBL; AE003803; AAF57871.1; -

DR HSSP; Q26998; 1BD3.

DR FlyBase; FBgn0034213; CG4798.

DR InterPro; IPR006083; PRK URK.

DR InterPro; IPR00764; Uridine_kin.

DR Pfam; PF00485; PRK; 1.

DR PRINTS; PRO0988; URIDINKINASE.

DR TIGRFAMs; TIGR00235; udk; 1.

KM Alternative splicing: Hypothetical protein.

FT VARSPIC 1 207 MISSING (IN SHORT ISOFORM).

SQ SEQUENCE 614 AA; 68704 MW; C81B6B3B9DB32A1 CRC64;

Alignment Scores:

Pred. No.: 6,4e-35 Length: 614

Score: 548.00 Matches: 116

Percent Similarity: 60.37% Conservative: 47

Best Local Similarity: 42.96% Mismatches: 61

Query Match: 18.47% Indels: 46

DB: 5 Gaps: 6

US-09-896-522-1 (1-1624) x Q9V810 (1-614)

QY 16 ACCTCGGCGCTGGGCGGCGC-----GCCGCGCGCGGGA----- 51

DB 132 ThrThrsertllegllyenglnThrThrAlaAsnProserGluCyslleleAla 151

QY 52 -----GGGCGCGCGCGCGGAGCCGATGCGCGGAGCGGAG 87

DB 152 AsnAsnArgThrIleTyrThrAlaGlyArgProProTyrTrpTyrAsnGlyGlnGln 171

QY 88 GCCGAGATGGCTTCGGCGGAGCGGAAGACTGCGAGAGCCCGCGGAGCCGACCT 147

DB 172 ValGlu----- 173

QY 148 CCGGACCGCGCGCCCTTCCTGATAGGGGTAGCGCGGCACTCCGCGGAAAGTGACC 207

DB 174 -----ProPheValIleGlyIleCysGlyGlySerAlaSerGlyThrThr 189

QY 208 GTGTGTGAGAAGATCATGAGTTGCTGGGACAGAACAGAGGTGAAACGCGCGGAG 267

DB 190 ValAlaGlnLysIleIleGluSerLeu-----AspAlaProTrp----- 202

QY 268 GTGTGATCATGAGCGGACAGAGGATTTCTACAGTCTGACGCGACAGAGGCGCAAG 327

DB 203 ValThrIleuSerMetAlaPcySPheTyrLysIleLeuAsnGlnLysGlnHISGln 222

QY 328 GCGCTTGAAGAGAGATGACATTTTGACCATCCAGATGCGCTTTATATGATTGAAGCAC 387

DB 223 AlaIleuIleAsnGlnLysIleuPheAspHisProAspAlaPheAspIleGluLeu 242

QY 388 AGGACTGTGAAGAACATCGTGGAGGCGAAGCGTGAAGGTGCGGACTATGATTGTG 447

DB 243 AspValIleuThrLysLeuLysGlnGlyArgLysValGlnValProValTyrAsnPheVal 262

QY 448 ACACTCAAGATTACAGAGACAGACAGGAGTGTCTACCTCGCGGACGTGTTGTTGAG 507

DB 263 ThrHisGlyArgLysSerGlnThrLysThrMetTyrGlyAlaAsnValIleIlePheGlu 282

QY 508 GCGATCTGTGTGTTCTACAGCGGAGATCCGGAATGTTCCACTGCGGCTTCTCGTG 567

DB 283 GlyIleuThrThrPheHisSerProGluValLeuLysLeuAsnAspMetLysIlePheVal 302

QY 568 GACACCGACTCCGACGTCAGGCTGTCTCGAAGAGTCTCCGAGAGCGTG---CGCGAGAG 624

DB 303 AspThrAspProAspIleArgLeuAlaArgValArgValArgAspIleSerGlnArgGly 322

QY 625 AGGAGACTGAGAGATTTGACGACGATACACCACTTGTGAAGCGCGCTTCGAGAG 684

DB 323 ArgAspLeuLysGlyValLeuLysGlnTyrLeuAsnMetValLysProSerTyrCysAsn 342

QY 685 TTTCGCTGCGGCAAGAAGATGCGCATGTGATCATCCACAGAGAGTGCACATATAG 744

DB 343 TyrIleAlaProThrMetAlaHisAlaAspIleIleValProArgGlyGlyAspAsnLys 362

QY 745 GTTGCGATCAACTGATCGTGGCGACATC 774

DB 363 ValAlaIleHisLeuIleValGlnHisVal 372

RESULT 11

Q8MOK4 PRELIMINARY; PRT; 626 AA.

AC Q8MOK4; (TREMBlrel. 22, Created)

DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)

DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)

DE L00359Sp.
GN CG4798.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Peiryoza;
NC Nematoda; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
XC NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkely,
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champ M., Chavez C., Dorsett V., Drenek D., Farfan D., Frise B.,
RA George R., Gonzalez M., Guanin H., Krommler B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunno J., Paclob J., Paragae V., Park S.,
RA Patel S., Phouenavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RL Celinker S.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY129436; AAU76178.1; -
DR FLYBase; FBgn0034213; CG4798.
DR InterPro; IPR006083; PRK_URK.
DR InterPro; IPR00764; Uridine_kin.
DR Pfam; PF00485; PRK_1.
DR PRINTS; PRO0088; URIDINKINASE.
DR TIGRFAMs; TIGR00235; udk: 1.
SQ SEQUENCE 626 AA; 70001 MW; 9DBEC0A7331AEF3 CRC64;

Alignment Scores:
Pred. No.: 6,43e-35 Length: 626
Score: 548.00 Matches: 116
Percent Similarity: 60.37% Conservative: 47
Best Local Similarity: 42.96% Mismatch: 61
Query Match: 18.47% Indels: 46
DB: 5 Gaps: 6

US-09-896-522-1 (1-1624) x Q8MQX4 (1-626)

QY	16	ACCTCGGCGCTGGGCGGCGGC-----GCCGCGCCCGGGGA-----	51
Dd	144	ThrThrSerIleGlyAsnGlnThrThrAlaIenProSerGluCyIleIleArgLa	163
QY	52	-----GGGGGGGGCGGGGAGCCCATGGCGGGGCGGAG	87
Dd	164	AsnAsnArgThrIleTyThrAlaGlyArgProPOTryTyrAsnCysValaGlyGln	183
QY	88	GCCGAGATGGCTTCGCGGGAGGCGAAACTCGAGAACCCTCGCGGAGCGCAACGT	147
Dd	184	ValGLu-----	185
QY	148	CSCGACCGAGCGCCTTCTTGATAGGGGTGACGGCGGCACCTCGACGGGAGTCAAC	207
Dd	186	-----ProheValIleGlyIleCysGlyGlySerIalaSerGlyLeuThrThr	201
QY	208	GTGTGTGAGAAGTCAATGAGATTGCTCGGGACAAGACGAGTGGAAACAGCGGCACGGAG	267
Dd	202	ValAlaGluLeuIleIleGluSerLeu-----AspValProTrp-----	214
QY	268	GTGTGATCTCGACGACGAGCATGTTTCAAGATCTCTGACGGGACGACAGAGGCAAG	327
Dd	215	ValThrLeuLeuSerMetCysPheTyTyrIleLeuAsnGluLeuGlnIleGluGln	234
QY	328	GCCCTGAAGAGACAGTCAATTTTGAACATCCATCCAGATGCTTGATATGATTTGATCAC	387
Dd	235	AlaLeuIleExonInLuTyrAsnPheAphisProdeltaPheAprIleGluLeuLeu	254
QY	388	AGGACCTGTGAAGAACATCGGTGAGAGGCAAAACGGTGGAGGTGGCGACCTATGATTTGTG	447
Dd	255	AspValIleuThrIleuLeuLeuGluGlyArgValaGluValProValTyrAsnPheVal	274
QY	448	ACACACTCAAGATTTAACAGAGCAACGCGGTCACTCCCTCGGAGAGTGATCTTTGAG	507
Dd	275	ThrHisGlyArgGluSerGlnThrIleThrMetTyrGlyAlaAsnValIleIlePheGlu	294
QY	508	GGCATCTTGATGTTTCAACGACGAGATCCGGACATGTTTCCATCTTGCGCTTCTGTG	567

Dd		225	GlylleleutInrPheHisSerProclivaileulySleuleuAAsmetLysIlePheVal	314
Oy		568	GACACCGACTCCGAGCTCAGGCCTGTCTTCGAGAATTCTTCCGGACCTG--CGCCGAGG	624
Dd		315	AspTrnAProAspIleArgLeunlaATgAArgLeuAArgApIlleGerlnArGly	334
Oy		625	AGGACCTCGAGACAGATTCTGACGCAATNACCACCTTGTAAGCCGCCCTTCGAGAG	684
Dd		335	ArgAapLeuLySGlyValleuLysgIntYrleuAemMetVallylserOserYrCyAsn	354
Oy		685	TTCTCCTCGCGACAAGATGATCCCATGATCATCCACGAGAGTGCACATPATG	744
Dd		335	TyrilleaAProThMeAlahisAlaspiellevalPronglYelalysaphanlys	374
Oy		745	GTTGCCATCAACTGATCGTCGAGACATC	774
Dd		375	ValAlatIleHleuIleValIdnlhlstval	384
 RESULT 12 Q8VYB2 PRELIMINARY; PRT; 466 AA.				
ID	Q8VYB2	01-MAR-2002 (TREMBLrel. 20, Last sequence update)		
AC	08VYB2	01-MAR-2002 (TREMBLrel. 23, Last annotation update)		
DT	01-MAR-2002 (TREMBLrel. 23, Last annotation update)			
DE	Puative uracil phosphoribosyltransferase.			
DN	ATI55810.			
OS	Arabidopsis thaliana (Mouse-ear cress).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;			
OC	eucotids II; Brassicales; Brassicaceae; Arabidopsie.			
OX	NCBI_TaxID=3702;			
XX	[1]			
RN	SEQUENCE FROM N.A.			
RA	Yamada K., Banb J., Chan M.M., Chang C.H., Chang E., Dale J.M.,			
RA	Deng J.M., Goldsmith A.D., Lee J.W., Onodera C.S., Quach H.L.,			
RA	Tang C., Toriumi M., Wu H.C., Yamamura Y., Yu G., Bowser L.,			
RA	Carrinai P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,			
RA	Kamiya A., Katlin-Neumann G., Kawai J., Kim C., Lam B., Iin J.,			
RA	Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,			
RA	Sakurai T., Satou M., Seki M., Shinn P., Southwick A., Shinzaki K.,			
RA	Davis R.W., Ecker J.R., Theologis A.,			
RT	"Full length cDNA of gene At1955810 (GI:15222778).";			
RL	Submitted (JAN-2002) to the EMBL/genbank/DBJ databases.			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RA	Yamada K., Chan M.M., Chang C.H., Dale J.M., Deng J.M., Heuan V.W.,			
RA	Lee J.M., Quach H.L., Tang C., Toriumi M., Wu H.C., Yu G., Bowser L.,			
RA	Carrinai P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,			
RA	Kamiya A., Katlin-Neumann G., Kawai J., Kim C., Lam B., Iin J.,			
RA	Miranda M., Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M.,			
RA	Seki M., Shinn P., Southwick A., Shinzaki K., Davis R.W., Ecker J.R.,			
RA	Theologis A.,			
RT	"Arabidopsis Open Reading Frame (ORF) Clones.";			
RL	Submitted (JUN-2002) to the EMBL/genbank/DBJ databases.			
DR	EMBL; AY072218; AAL60039.1; -			
DR	EMBL; AY122946; AAM64749.1; -			
DR	InterPro; IPR006083; PRK URK.			
DR	InterPro; IPR00764; Uridine_kin.			
DR	Pfam; PF00485; PRK_1.			
DR	PRINTS; PRO0968; URIDINKINSE.			
DR	TIGRFAMS; TIGR00235; udk; 1.			
KW	Glycosyltransferase; Transferase.			
SQ	SEQUENCE 466 AA; 52443 MW; 3FP6783CB89727F7 CRC64;			
 Alignment Scores:				
Pred. No.:	1.23e-34	Length:	466	
Score:	544.00	Matches:	116	
Percent Similarity:	63.16%	Conservative:	52	
Best Local Similarity:	43.61%	Mismatches:	78	
Query Match:	18.34%	Gaps:	20	
DB:	10	Indels:	8	

Db	350	CysCysLysGlyIleIleValIleGlyLys-----IleuLeuLeH 362
Oy	1027	CCCCAGGGGA-----GTTTGCACGAGGCTTC 105
Db	362	sargGlugIysapaemgIyGInGInValCysValLeuSerLeuIleIleThsSerProAs 382
Oy	1057	CTCACTCAGACAGTGGAAAC-----TCACATGN 108
Db	382	nTyrlreuleuThrtrHtrHAsnGlyThrhIsGInleuIleTyrgIuLysLeuProSerAspII 402
Oy	1084	GTCACTCAGACTCACTCACTGCTG 1105
Db	402	eserGluArghIsValLeuLeu 409
RESULT 15		
O9LG32		PRELIMINARY; PRT; 542 AA.
AC	O9LG32;	
DT	01-OCT-2000 (TREMBLrel. 15, Created)	
DT	01-OCT-2000 (TREMBLrel. 15, Last sequence update)	
DT	01-OCT-2000 (TREMBLrel. 22, Last annotation update)	
DE	FL416.5.	
OS	Arabidopsis thaliana (Mouse-ear cress).	
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;	
OC	eurosid II; Brassicales; Brassicaceae; Arabidopsie.	
OX	NCBI_TaxID=3702;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RA	Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C., Khan S.,	
RA	Kim C., Altafi H., Bei Q., Chin C., Chio J., Choi E., Conn L.,	
RA	Conway A., Gonzales A., Hansen N., Howing B., Koo T., Lam B., Lee J.,	
RA	Lenz C., Li J., Liu A., Liu K., Liu S., Mukharshy N., Nguyen M.,	
RA	Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thavert A.,	
RA	Toriuna M., Vaysberg M., Yu G., Federspiel N.A., Theologis A.,	
RA	Ecker J.R.;	
RT	"Genomic sequence for Arabidopsis thaliana BAC FL416 from chromosome	
RT	1."	
RL	Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.	
RN	[2]	
RP	SEQUENCE FROM N.A.	
RA	Ecker J.R.;	
RL	Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.	
RN	[3]	
RP	SEQUENCE FROM N.A.	
RA	Ecker J.R.;	
RL	Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.	
RN	[4]	
RP	SEQUENCE FROM N.A.	
RA	Ecker J.R.;	
RL	Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.	
RN	[5]	
RP	SEQUENCE FROM N.A.	
RA	Ecker J.R.;	
RL	Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.	
RN	[6]	
RP	SEQUENCE FROM N.A.	
RA	Cheuk R., Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C.,	
RA	Khan S., Kim C., Altafi H., Bei B., Chin C., Chio J., Choi E.,	
RA	Conn L., Conway A., Gonzalez A., Hansen N., Howing B., Koo T., Lam B.,	
RA	Lee J., Lenz C., Li J., Liu A., Liu J., Liu S., Mukharshy N.,	
RA	Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,	
RA	Thavert A., Toriuna M., Vaysberg M., Yu G., Davis R., Federspiel N.,	
RA	Theologis A., Ecker J.;	
RL	Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.	
RN	[7]	
RP	SEQUENCE FROM N.A.	
RA	EMBL; AC002304; AAF79310.1; ..	
DR	HSSP; Q26998; 1BD3.	
DR	InterPro; IPR006083; PRK URK.	
DR	InterPro; IPR000764; Uridine_kin.	
DR	Pfam; PF00485; PRK; 1.	
DR	PRINTS; PRO0988; URIDINKINASE.	
DR	SEQUENCE 542 AA; 61022 MW; 35427303EBF73942 CRC64;	

Alignment Scores:	
Pred. No.:	1 02e-31
Score:	507.50
Percent Similarity:	56.19%
Best Local Similarity:	58.80%
Query Match:	17.10%
DB:	10
Gaps:	10
Length:	542
Matches:	116
Conservative:	52
Mismatches:	78
Indels:	53
Gaps:	10

US-09-896-522-1 (1-1624) X Q9LG32 (1-542)

QY 124 AGCCCCGCGCGAG-----GCCGACCGTCCGCACCAAGCGGCCCTTCTGATA--- 171

Db 29 SerAsnArgProGluGlnMetAlaGluGluGluGluHisGlyGlnProPheValIleGly 48

172 -----GGGGTGAGCGGCGCACTGCCAGCGGGAAGTCGACCCGTGTGTGAGAAGATC 222

Db 49 GluSerArgGlyValAlaGlyGlyAlaAlaSerGlyLysThrThrValCysAspMetIle 68

223 ATGAGTTGCTGGACAGACGAGGTGGAACAGCGCAGCGGAAGTGTCTCCTGAGC 282

Db 69 MetGInLeu-----HisAspGInArgAlaValValAsn 81

QY 283 CAGGACAGGTTCTACAGGTCCTGACGCGCAGAGCAGAAGGCCAAGGCTTGAAAGGACAG 342

Db 82 GlnAspSerPheTyrHisAsnValAsnGluValGluLeuValArgVal-----HisAsp 99

343 TACAATTTGACCATCCAGATGCCCTTGATAATGATTGTATGCACAGGACTCTGAAGAAC 402

Db 100 TyrAsnPh~~e~~AspHisProAspAlaPh~~e~~AspThrGluGlnLeuLeuSerSerMetGluLys 119

403 ATCGTGAGGGCAAAACGGTGGAGGTGCCGACCTATGATTTTGTGACACACTCAAG-- 459

Db 120 LeuArgLysGlyGlnAlaValAspIleProAsnTyrAspPheLysSerTyrLysAsnAsn 139

460 ---TTACCAGAGACCACGGTGGTCTACCCCTGCCGACGTGTTCTGTTGAGGCATCTTG 516

Db 140 ValPheProPArgarg---ValAsnProSerAspValIleIleLeuGluGlyIleLeu 158

517 GTGTTCTACAGCCAGGAGATCCGGACATGTTCCACCCTGCGCCTCTTCGTGACACC--- 573

Db 159 ILEPHEHISAPRROALGVALARGAPLEUMETASMETLYSILEPHEVALASPRALAGLY 178

QY 573 ----- 573

Db 179 LeuSerHisThrLysProValAsnThrTyrValValLysSerValAlaTyrMetArgArg 198

574 -----GACTCCGACGTCAGGCTGTCTCGAAGAGTTCTTCCGG 609

Db 199 CystHrcysIecCystHrcHisGluAspAlaAspValArgLeuAlaArgArgIleLysArg 218

610 GAC---GTGGCGCGAGGGGACCTGGAGCAGATTCTGACGCAGTACACCACCTTCGTG 666

Db 219 AsptHrValGluLysGlyArgAspIleAlaThrValLeuAspGlnTyrSerLysPheVal 238

667 AAGCCGGCCTTCGAGGAGTCTGCCTGCCGACAAAGAAGTATGCCGATGTGATCATCCCA 726

Db 239 LysProAlaPheGluAspPheIleLeuProThrLysLysTyrAlaAspIleIleIlePro 258

QY 727 CGAGGAGTGCACATATGGTTGCCATCAACCTGATCGTGCAGCACATCCAGGACATTCTG 786

Db 259 ArgGlyGlyAspAsnHisValAlaIleAspLeuIleValGlnHisIleHisThrLeu 278

QY 787 --AATGGTACATCTGCCAATGGCACCGAGGGTCCAAATGGGGGAGCTACAAGCGG 843

Db 279 GLYGLNHISAspleucyslyleTyr-----ProAsnleuTyrValIleGlnSer 295

844 ACCTTTCTGAGCCAGGGACCACTGGGATGCTGACCCTCTGGCAACGGTCACAT 900

Db 296 ThrPheGlnIleArgGlyMetHisThrLeuIleArgAspSerLysThrThrLysHis 314

Search completed: November 25, 2003, 07:38:50
Job time : 125.819 secs